THESIS

UNDERSTANDING RATEE PREFERENCE FOR FEEDBACK FORM THROUGH RATING FORMAT AND EXPECTATION CLARITY

Submitted by

Jaclyn M. Menendez

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2014

Master's Committee:

Advisor: Jeanette Cleveland

Kurt Kraiger
Karen Rambo-Hernandez
Bryan Dik

ABSTRACT

UNDERSTANDING RATEE PREFERENCE FOR FEEDBACK FORM
THROUGH RATING FORMAT AND EXPECTATION CLARITY

Performance feedback is one of the most highly researched areas in industrial/organizational psychology, but it remains far from being fully understood. Ratee attitudes towards their feedback forms can shape their attitudes about the feedback content in general. This study investigates the impact of two feedback form features, rating format and performance expectation clarity, on two attitudinal outcomes of satisfaction and acceptance. Perceptions of fairness were hypothesized to mediate all relationships. Results showed a statistically significant relationship between clarity of expectations and ratee satisfaction with their feedback form. All other relationships were found to be non-significant. This research demonstrated the importance of performance expectation clarity on ratee satisfaction with feedback. The results are discussed in relation to future research and implications for behavioral change.

TABLE OF CONTENTS

INTRODUCTION AND LITERATURE REVIEW

The appraisal of employee performance is central to human resource management (Boswell & Boudreau, 2002; Kuvaas, 2006). The feedback provided through these appraisals is an important source of information for employee development, improvement, and motivation (Fletcher, 2001). Managers use feedback to communicate with employees about individual performance, developmental opportunities, and strategic goals (Bacal, 2004; Dusterhoff, Cunningham, & McGregor, 2013). The most common formal medium for delivering this type of feedback is in written form (Alvero & Bucklin, 2001). Successful strategic management often relies on these written performance feedback forms to deliver the information that links an individual's capabilities to the organization's goals and mission (Dusterhoff et al., 2013).

Despite the benefits that written performance feedback can provide to an organization, many employees report dissatisfaction with their performance appraisal and feedback process (Fletcher, 1997). This dissatisfaction stems from a number of obstacles and antecedents, including conflict between supervisors and employees over negative feedback (Lawler, 1994), the tendency to ignore feedback that is incongruent with one's self-perception (Keeping & Levy, 2000), and issues with the feedback form itself are a few examples (Posthuma & Campion, 2008). The current research addressed two issues associated with the feedback process: the feedback form's rating format; and the clarification of performance expectations.

Employee attitudes towards feedback can range from positive (e.g. accepting the feedback and having intentions to use it), to neutral/dismissive (e.g, ignoring the feedback or forgetting about it), to negative (e.g. dissatisfaction, feelings of unfairness, etc.). Negative attitudes have potentially serious repercussions, as negative attitudes regarding performance

feedback have been associated with more developmental harm than good (Alvero et al., 2001; Keeping et al., 2000). When employees do not view the feedback as useful or accurate, they are not likely to change their behavior in response to it (Keeping et al., 2000). Therefore, it benefits organizations to craft a feedback form that presents information in an appealing and effective manner.

The proposed research explores two specific aspects of written feedback forms: rating format and clarity of expectations. These factors have been chosen based on past research and theory that identify these factors as potential influences on employee feedback attitudes (Atwater & Brett, 2006; Kline & Sulsky, 2009). The goal of the proposed research is to examine each of these factors in depth, and evaluate their effect on the specific employee attitudes of acceptance, satisfaction, and fairness. The two independent variables will also be considered together to assess a possible interaction between feedback rating format and expectation clarity on employee attitudes towards feedback forms.

Historically, the methods for improving feedback systems focused upon improving the format and measurement accuracy (Kline et al., 2009). A landmark review by Landy and Farr (1980) shifted the focus of appraisal research from largely construct measurement issues to rater training, accuracy, and the context of rating. In order to establish effective performance feedback as a whole, more recent research suggests that rater and ratee perception of and reactions to feedback are critical (Murphy & Cleveland, 1995). The current research examines both a key process variable (e.g. clarity of performance expectations) as well as a key content variable (e.g. feedback rating format) and investigates how these factors are associated with ratee satisfaction and acceptance.

The current research expands on the theories of ratee attitudes and preferences, in addition to incorporating a strong practical approach that should appeal to organizations. By determining if slight manipulations to the rating format and clarity of performance feedback forms can meaningfully influence employee attitudes, organizations can invest in these changes accordingly. With the rising popularity of costly performance appraisal software and outsourcing (Robb, 2008), it is clear that many organizations feel ill-equipped to handle performance appraisals and feedback systems on their own (Fletcher, 1997; Robb, 2008). Perhaps with more straightforward and easily applicable suggestions, companies will become better equipped to use their own feedback forms, as well as making sure that they are developed soundly.

In order to understand the relationships proposed by this research, a brief history of feedback research will be presented with a focus on feedback forms. The independent variables of feedback rating format and clarity of expectations will then be explored in detail, focusing on their influence on a ratee's acceptance and satisfaction with feedback forms. The potential mediating effect of perceived fairness and organizational justice on these relationships will also be explored.

Defining Terms

When attempting to untangle the classic and current issues of feedback in the workplace, it is important to have a clear understanding of common terms and definitions. Performance feedback and their forms are most often associated with performance appraisals. Performance appraisals are a "general heading for a variety of activities through which organizations seek to assess employees and develop their competence, enhance performance and distribute rewards" (Fletcher, 2001, pp. 474). Performance appraisals are also commonly used to inform and support employment decisions, such as promotions or salary negotiations (DeNisi & Pritchard, 2006).

In slight contrast, performance feedback is not necessarily tied to pay, promotion, or anything other than information about past performance (Balcazar, Hopkins, & Suarez, 1986; Prue & Fairbank, 1981). Performance feedback is information about an individual's adequacy in their past behavior and/or performance, usually with the goal of improving future performance (Ilgen, Fisher, & Taylor, 1979). More specifically, performance feedback is a systematic way to provide information about an employee's performance and development (DeNisi & Pritchard, 2006). Performance feedback can also be understood as "...actions taken by (an) external agent(s) to provide information regarding some aspect(s) of one's task performance" (Kluger & DeNisi, 1996, pp. 255). This feedback is typically based on established objectives and organizational criteria, and can be used as a calibration tool for employees to understand what is expected of them and how well they are meeting those expectations (Muchinsky, 2012). The medium(s) used to deliver performance feedback vary by organization, with written feedback as the most common type (Alvero et al., 2001). With changing technologies and delivery methods, it may technically be more appropriate to interpret the term "written" feedback as "read" feedback, as written feedback can also refer to typed feedback.

Background of Performance Feedback Research

In order to provide context to the relationship being investigated, a brief background on performance feedback research is necessary. The first studies on feedback and feedback manipulations date back to the 1900's (Arps; 1920; Gates, 1922; Kluger et al., 1996). These earliest studies focused mainly on how feedback was linked to changes in performance. However, these studies had a number of methodological and measurement limitations including small sample sizes, questionable interpretations, confirmation bias, and a lack of operational definitions that resulted in misleading results (Kluger et al., 1996). Most importantly, the lack of

an operational definition for "performance feedback" precluded many of these studies from being generalizable, as they were often measuring different constructs. Studies that claimed to simply be examining the effect of performance feedback on behavior changes would also include other variables (e.g. incentives to do better, punishments for poor performance, the ability to practice) without insight or clarification as to how these additional variables might be moderating the relationship between knowledge of results and feedback effectiveness (Crawley, 1926; Gates, 1917; Kluger et al., 1996; Thorndike, 1927).

Current Issues and State of the Field

Performance feedback remains a vital and thriving research area to date. Alvero, Bucklin & Austin (2001) published the most recent literature review of performance feedback in organizational settings as a whole, spanning the field from 1985-1998. They detailed findings that highlighted common practices in contrast to best practices. For example, most of the feedback studies used supervisors to provide ratings, even though studies that used both supervisors and researchers yielded the most consistently effective feedback. Additionally, most feedback was delivered by written report (Calpin, Edelstein, & Redmon, 1988; Fox & Sultzer-Azaroff, 1989; Gaetani, Hoxeng, & Austin, 1985; Petty, Singleton, & Connell, 1992; Siero, Boon, Kok et al., 1989), yet the most consistently effective feedback was delivered as a combination of written, verbal, and graphical information (Babcock, Sultzer-Azaroff, & Scibak, 1992; DeVries, Burnette, & Redmon, 1991; Hawkins, Burgio, Lanford, & Engel, 1992; Nordstrom, Lorenzi, & Delquadri, 1988; Porterfield, Evans, & Blunden, 1985).

Multi-source feedback. Another strong focus within performance appraisal research is the effects and effectiveness of multi-source feedback (Gregarus, Ford & Brutus, 2003). London, Smither, & Riley (2002) conducted the most recent meta-analysis on this specialized topic.

Multi-source feedback is most commonly referred to as 360 degree feedback, and it pulls information from various resources (coworkers, subordinates, managers) to ostensibly provide a more complete and informed feedback experience. The potential issues facing this type of feedback are a popular research topic; namely, inter-rater disagreement (Gregarus et al., 2003), how ratee characteristics can influence its effectiveness (Kulas & Finklestein, 2007), and the dangers of replacing traditional performance appraisals entirely with 360 assessments (Toegel & Conger, 2003). However, some research has also linked multi-source feedback to beneficial outcomes, which include an increased positive perception of the organizational climate (Mamatoglu & Baysal, 2008). The majority of these studies conclude that multi-source feedback should not be universally embraced without seriously considering the drawbacks, such as the overload of information, the likelihood of disagreement between groups, and the statistically tenuous relationships with improved performance.

Although the proposed research focuses on general performance feedback rather than multisource feedback, issues of information overload and disagreement between groups are related to the variables of rating format and clarity of expectations. As will be discussed, the results of this research on feedback forms have potential applicability to both single and multi source feedback situations.

Ratee preferences and attitudes. One trend in performance feedback research, and the focus of the present study, is ratee attitudes towards performance feedback. It is widely accepted that the most effective feedback is reliable, valid, practical, and accepted by users (Ilgen & Barnes, 1993, Latham & Wexley, 1981; Thorndike, 1949). Feedback that fulfills these qualities tends to lead to higher rates of behavioral change that align individual performance with the organization's goals (Ilgen et al., 1993). Despite theoretical linkages of feedback success to

general job satisfaction, ratee attitudes towards feedback have received little attention in actual research (Keeping et al., 2000; Pichler, 2012; Pettijohn, Pettijohn, & d'Amico, 2001). In the past, most research has been conducted on how to minimize rater error and increase accountability for the raters who are conducting the performance appraisals and the managers that are delivering performance feedback (Bazerman, Beekum, & Schoorman, 1982; Kingstrom & Mainstone, 1985; Mero, Guidice, & Brownlee, 2007; Murphy & Cleveland, 1995), with less focus on ratee's reactions and attitudes towards feedback. Murphy and Cleveland (1995) coined the phrase "neglected criteria" (pp. 310) to refer to this issue, and it has continued to be an under-researched area of feedback (Cawley, Keeping, & Levy, 1998; Culbertson, Henning, & Payne, 2013). The study of ratee attitudes towards performance feedback and feedback forms therefore fills an important science-practitioner gap in I/O psychology (Keeping et al., 2000).

Ratee attitudes towards feedback have gained attention in recent years because of the immediate and long-term effect these attitudes can have on the ratee's likelihood of using the feedback (Jawahar, 2006). The majority of research on ratee reactions to feedback has focused on attitudes towards the actual feedback, and how these attitudes influence the likelihood of using feedback to make behavioral changes to performance. The present study focuses on attitudes towards the specific feedback form as an indicator of their likelihood to use the feedback for behavioral change in the future.

Attitude-Behavior Theory

The way that feedback is presented has an indirect influence on the ratee's likelihood of using that feedback for behavioral change (Matsumara & Hann, 2004). The presentation of the feedback form contributes to ratee attitudes towards the feedback in general, but the extent of this relationship is not well understood (Matsumara et al., 2004). Using Ajzen's (1991) theory of

planned behavior, the relationship between attitudes toward feedback forms and behavioral changes will be explored in depth, with an understanding that measuring the outcome of ratee preferences and attitudes should greatly contribute to building a more effective feedback system overall.

Preferences linked to attitudes. Ratee preferences for feedback forms have been linked to their attitudes about the feedback in general (Gosselin, Werner, & Halle, 1997, Matsumara et al., 2004). When preferences for a feedback type match the received feedback type, ratee attitudes towards the feedback in general are more positive (Gosselin et al., 1997). The specific attitudes of acceptance and satisfaction with feedback tend to be influenced when preferences are matched (Manshor & Kamalanabhan, 2000). The present research measures both ratee preferences and ratee attitudes towards feedback type, with an understanding that the two should yield similar results.

Attitudes linked to intentions to change. An attitude is "a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor" (Eagly & Chaiken, 1993, pp. 1). Attitudes towards an object are believed to influence intentions to behave a certain way towards that object (Ajzen, 1991). Therefore, attitudes towards feedback influence one's intention to use that feedback to change their behavior (Hedge & Borman, 1995; Keeping et al., 2000; Murphy & Cleveland, 1995). Specifically, attitudes of acceptance and satisfaction are crucial if the performance feedback is to be effective (Murphy & Cleveland, 1995). If the ratee reacts poorly to the feedback, the utility of the feedback system is compromised (Jawahar, 2006; Murphy & Cleveland, 1995). As the feedback form contributes to attitudes towards the feedback in general, it stands to reason that the benefits of a feedback system also rest on the recipient's attitude toward the general presentation of the feedback (Matsumara et al., 2004).

Intentions linked to behavior. Behavioral intentions are often used in research as a proxy for actual behavior, because these intentions represent the degree of motivation we have to perform a behavior (Ajzen, 1991). However, there is more to performing a behavior than simply having the motivation to do so. Those intentions are influenced by attitudes towards the behavior, subjective norms, and perceived behavioral control (Ajzen, 1991). Attitudes, as discussed above, can be understood as a positive or negative evaluation of performing a certain behavior. Subjective norms address the social implications of engaging in a certain behavior. Behavioral control addresses the perceived ability and opportunity for a person to engage in the intended behavior. This theory of planned behavior states that intentions are most likely to predict behavior when all three of these conditions are met (Ajzen, 1991; Ajzen & Madden, 1986). See Figure 1 below.

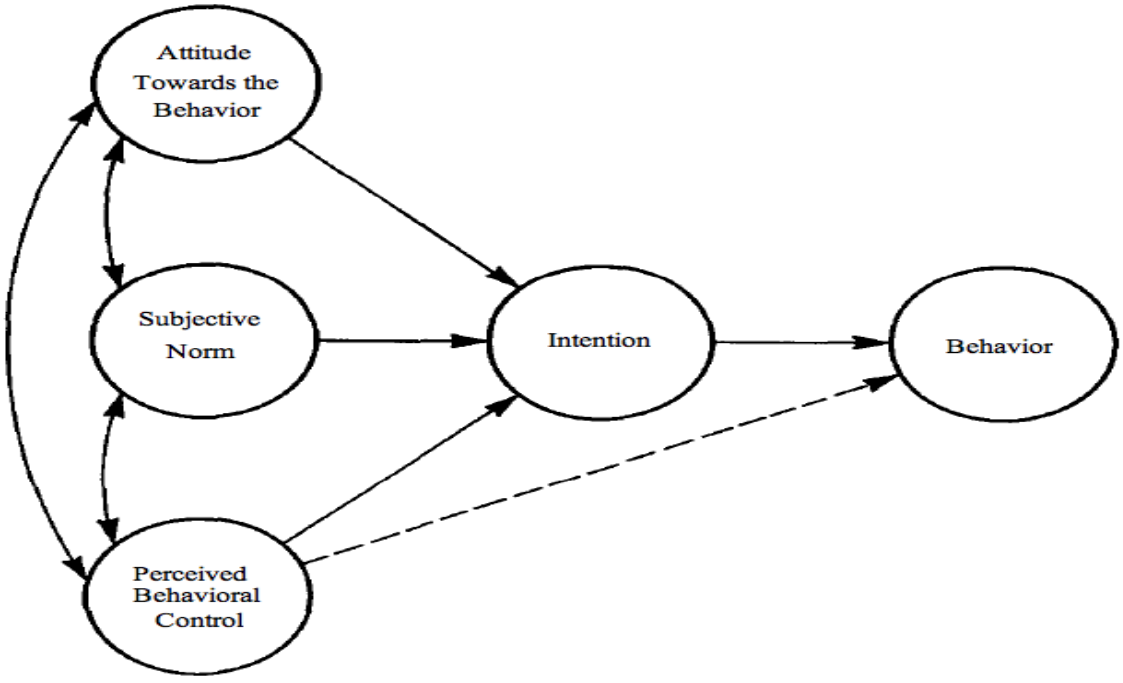


Figure 1. Theory of Planned Behavior.

To summarize, attitudes are an important piece of behavioral intentions, and research has shown that intentions to engage in a behavior are predictive of actual behavior. The present study hypothesizes that attitudes towards feedback forms are theoretically useful in determining the likelihood of a ratee using that feedback to make behavioral changes. Specifically, alterations to the feedback rating format and expectation clarity may influence a participant's satisfaction and acceptance of the feedback form. The importance of these specific attitudes will be explored next, followed by the proposed variables of influence.

Satisfaction and Acceptance

After reviewing the process that links a ratee's preferences and attitudes to their behavioral outcomes, the specific attitudes of satisfaction and acceptance require further explanation. It is not uncommon for feedback studies to combine these two attitudes, and many similar attitudes, into the overall construct of "appraisal reactions". In scales that measure the attitudes of feedback recipients, satisfaction and acceptance are often measured by the same question or implied to be synonymous. For example, a recent meta-analysis by Pichler (2012) examined "appraisal reactions" which consisted of accuracy, fairness, utility, and satisfaction with the feedback process. Though past research supports a correlation between satisfaction and acceptance (Jawahar, 2006), empirical investigations have shown that these constructs are not identical and should be measured separately (Keeping et al., 2000). This research therefore examined satisfaction and acceptance as two separate outcomes.

Feedback Acceptance. Feedback acceptance is a broad term, and can mean many things depending on the context. Russell and Goode (1988) define feedback acceptance as how *accurate* an individual finds their feedback, and Ilgen et al.'s (1979) landmark research on the topic similarly defined feedback acceptance as "the recipient's belief that the feedback is an

accurate portrayal of his or her performance" (pp. 356). Ilgen et al. (1979) makes the important distinction that this perception is not necessarily rooted in reality; whether or not a ratee is accurate in their self-perception is irrelevant here. For the present study, a slight modification of these terms was necessary to define feedback *form* acceptance. *Feedback form acceptance* can be understood as a ratee's perception of the form's accuracy in presenting information.

Satisfaction. Satisfaction is believed to be the most influential predictor of feedback effectiveness (Culbertson et al., 2013; Jawahar, 2006), and it is certainly the most frequently measured attitude in the feedback process (Giles & Mossholder, 1990; Keeping et al., 000;). Satisfaction with feedback has been one of the most studied outcomes of the performance feedback process because of its established connection with increased productivity, motivation, and commitment to the organization (Cawley et al., 1998; Ilgen et al. 1979). Jawahar (2006) found that employee satisfaction with performance feedback influenced future performance after controlling for past performance, job satisfaction, and manager satisfaction. For the purpose of this research, we can generalize *satisfaction with feedback* to also reflect *satisfaction with the feedback form*. Therefore, *satisfaction with feedback form* can be understood as the extent to which a person is satisfied with the feedback form.

Feedback Rating Format

In trying to maximize the attitudes of acceptance and satisfaction towards a feedback form, research has shown that small changes can be effective. Seemingly slight alterations to feedback presentation can influence attitudes towards the feedback form, and the feedback in general (Atwater et al., 2006). One feedback form feature that can be easily manipulated is the rating format. *Rating format* refers to the method used to convey performance ratings, and for the purpose of this research, can be understood as numerically based (e.g. "you are a 3.5 out of 5 on

this task") or text based (e.g. "you are performing below average on this task"**)** (Atwater et al., 2006). Feedback rating format can take many other forms, such as narrative or graphical, but the present research examines the most commonly type of distinction.

The importance of rating format has received attention in the past, though not without criticism. For example, Atwater et al. (2006) examined how employees' performance had changed nine months after receiving one of two feedback forms. One form contained number-based ratings, and the other was text based. They hypothesized that text-based rating options would be more acceptable to participants, particularly when negative feedback was given. They rationalized that text was less threatening, and less likely to result in normative comparisons than numerical ratings. This was based on previous research that found numerical rating formats prompted participants to compare themselves more to their peers, which diverted attention away from the task at hand and thus resulted in lowered performance (Ilgen & Davis, 2000). This effect was found to increase in the presence of negative feedback (Bobko & Colella, 1994).

Atwater et al. (2006) found that, contrary to their hypothesis, participants significantly preferred receiving feedback that was *numerically* based rather than text based. Participants specifically identified that the numerical rating format was easier to understand, more concise, and seemed more specific (Atwater et al., 2006). The study also concluded that the numerically based rating format actually encouraged participants to focus on performance, while the ambiguity of the text based rating format prompted people to compare their ratings with peers.

While these results were contrary to their original hypothesis, the present study supports the findings of Atwater et al. (2006). The study's qualitative component provided clear support that these results were due to real differences in preference, and raised the possibility that numerical rating formats were inherently more appealing than text rating formats. It is also

important to note that this study was the first of its kind; no previous research had differentiated ratee reactions based solely on feedback's text/numerical format (Atwater et al., 2006; Feys, Anseel, & Wille, 2011). Due to the exploratory nature of this research question, coupled with the supporting qualitative evidence, the results of this experiment indicate that numerical rating formats may be the most appealing type to use within feedback forms.

Following the publication of these results, critics pointed to the possibility that differences in reactions were actually due to one form containing more information than the other. Therefore, participants may have actually chosen the more informative of the forms rather than basing their choice on rating format alone (Anseel & Lievens, 2009). As this study was one of the first attempts to differentiate ratee reactions based solely on feedback's rating format, it has been difficult to assess the accuracy of these criticisms (Atwater et al., 2006; Feys, Anseel, & Wille, 2011). Subsequent studies have altered other detail-oriented aspects of feedback forms, such as type and amount of information given (Feys et al., 2011), but no direct replication of Atwater et al.'s study has been published. The present study aims to not only support their findings via replication, but also improve upon the past study's experimental conditions. Specifically, I address their limitation by ensuring that both rating formats are identical in the amount of information provided.

To conclude, the immediately-formed attitude that a participant has to their feedback determines their long-term performance changes (Atwater et al., 2006). Feedback rating format has shown importance in shaping ratee attitudes towards their feedback (Atwater et al., 2006). An immediately favorable reaction towards a feedback form's presentation can predict positive attitudes towards the feedback in general, as well as changes in behavior at a later time (Atwater et al., 2006). Conversely, people with negative attitudes towards the feedback form tend to

struggle with accepting the feedback (Atwater et al., 2006). This stresses the importance of creating feedback forms that will immediately appeal to the receiver. I hypothesize that numerically based feedback will accomplish that goal best.

H1: When the information is identical, participants will prefer numerically based feedback to text based feedback.

Clarity of Performance Expectations

In addition to feedback rating format, the expectations for performance must be clearly represented in a feedback form. As a goal of feedback is to inform the employee about his or her performance, the organization's definition of "performance" must be clear. *Clarity of performance expectations* can be understood as clear and comprehensive information regarding how an organization measures the performance of its employees.

Clarity of expectations is important because every organization has their own standards for how to measure performance. Performance can be evaluated based on task, context, counterproductive work behaviors, adaptive performance, or a mixture of these and other components (Borman & Motowidlo, 1993; Murphy & Cleveland, 1995). Even within these dimensions, there can be varying levels of clarity of the specific behaviors that are being evaluated in terms of frequency, importance, and many other factors.

Without a clear explanation of performance expectations, the face validity of performance ratings can be viewed as questionable by ratees (Kline et al., 2009). Several job attitudes may be negatively impacted by unclear expectations as well (Greenberg & Colquitt, 2013). If employees are not adequately informed of the organization's expectations or evaluation

standards, they will likely experience lowered satisfaction with and acceptance of feedback. Low satisfaction with feedback processes may ultimately result in lowered overall job satisfaction (Behrman, Bigoness, & Perrault, 1982; Brown, Hyatt & Benson, 2009).

In addition to attitudinal changes, ratees may also make misguided behavioral changes in response to unclear feedback. Employees may try to extrapolate the feedback's message in order to meet the unclear performance expectations, but the employees' perceptions of successful behavior changes may be misaligned with what the organization intended (Ilgen et al., 1979). For example, an organization might consider "being at work on time" to be an important part of job performance. However, in performance feedback sessions, a measurement of timeliness is never directly explained or detailed. A habitually late employee that then receives negative feedback on their performance may be confused or frustrated, as the standards were never clearly communicated and they are unsure how to improve. Without knowing the specific behavior that led to the negative feedback, the employee may make irrelevant changes to their behavior in a misguided attempt to improve their performance (e.g. staying late at the office, volunteering for extra assignments, etc). When the employee continues to receive the same vague negative feedback based on vague performance standards despite their efforts to improve, they may become frustrated and dissatisfied with the performance feedback process altogether. Therefore, unclear expectations may cause ratees to make changes to their performance and behavior after the performance feedback process, which is not advisable and will likely decrease their perceptions of fairness or satisfaction with the received feedback. (Ilgen et al., 1979).

The importance of expectation clarity has roots in goal setting theory (Latham &Yukl, 1975; Locke, 1968). This theory posits that the clarity and the difficulty of a goal can influence the likelihood of reaching that goal. Goals that are specific and ambitious tend to result in the

most success, relative to goals that are vague, too easy, or unrealistic (Latham & Yukl, 1975). This theory can be applied to performance feedback, in that behavioral changes based on feedback will be most successful when a clear and specific goal has been established (Locke & Latham, 1990). However, clear and specific goals to achieve a desired performance level cannot be attained if the organization provides vague expectations or unclear information about the range of possible performance levels (Katz & Kahn, 1978). Therefore, feedback forms that are low in clarity of expectations do not provide ratees with the necessary framework to develop productive or specific goals.

For these reasons, I hypothesize that feedback forms containing clear performance expectations will receive more favorable reactions than feedback forms that are vague or unclear. People will likely feel more satisfied with a form that communicates clear expectations, and will be more willing to accept these forms as accurate feedback tools. Feedback forms with clear expectations will make it easier to form accurate and useful goals to meet those expectations.

H2: Participants will prefer feedback that is high in clarity of expectations.

Methods of improving the clarity of expectations are varied. In general, high clarity of feedback standards consists of highlighting strengths, developmental opportunities, and specific recommendations for improvement (Denton, Madden, Roberts, & Rowe, 2008). Expectation clarity has also been measured by using clear indications of the relevant performance dimensions, the behaviors associated with different performance levels, and the possible range of performance levels (Cardy & Keefe, 1994; Kline et al., 2009). This research incorporates Cardy et al.'s (1994) definition when defining the measure of expectation clarity, which will be explained in greater detail within the Measures section.

Interaction between clarity and rating format. The variables of expectation clarity and feedback rating format are expected to influence ratee preference both separately and together. A significant interaction between these two variables is predicted. Specifically, feedback rating format will be most influential when the clarity of expectations is low. As clear standards are needed by ratees to understand expectations and subsequent feedback (Kline et al., 2009), in the absence of clear direction it is likely that a numerically based feedback rating format will continue to be the preference of ratees (Atwater et al., 2006). For measures of high feedback clarity, the feedback rating format will influence a ratee's preference to a lesser and likely non-significant degree. This proposed interaction has not been studied in past research and is exploratory in nature, though it draws conceptually from the research conducted on each individual variable.

> *H3*: *There will be an interaction between rating format and expectation clarity, such that participants will prefer numerically based feedback forms in low clarity conditions, and this preference will be less in high clarity conditions.*

Justice and Fairness

Related to the attitudes of satisfaction and acceptance, organizational justice and fairness perceptions play an important role within performance feedback (Bretz, et al., 1992; Keeping et al., 2000; Smither, 1998). *Organizational justice* can be understood as perceptions of workplace fairness (Byrne, 2005). The importance of this attitude spans research and practice, with managers reporting fairness as the most important issue they face in the performance appraisal and feedback process (Bretz et al., 1992). Feelings of unfairness in a feedback environment are closely tied to a lack of acceptance or buy-in of feedback. The result of these negative reactions

can range from feedback having no effect on the ratee, to a strong negative effect that results in decreased performance and counterproductive behavioral change (Cardy & Dobbins, 1994). Conversely, strong and positive attitudes of fairness are positively tied with organizational commitment, lowered job-related stress, and high job satisfaction (Byrne, 2005; Cohen-Charash & Spector, 2001). These links can be generalized to feedback research, in that satisfaction with a form may be a result of the perceived fairness of the form.

Feedback acceptance is also closely tied with perceived fairness of the feedback. The relationship is thought to be cyclical, as informative feedback is a key aspect of fairness perceptions within organizations (Anseel et al., 2009), and fairness affects a wide range of on-the-job behaviors and attitudes, including acceptance of feedback. Clear feedback, particularly when it focuses on specific behaviors, increases feelings of fairness, which in turn leads to higher rates of accuracy perceptions (Van Vianen, Taris, Scholten, & Schinkel, 2004). Therefore, this research posits that when feedback forms appear fair, ratee's acceptance of the form would increase as a result.

Early fairness research in terms of performance feedback was general, and mainly examined perceptions of fairness towards the system in general (Keeping et al., 2000; Landy & Farr, 1980). Since then, we have learned to separate organizational fairness into distinct categories, which now include distributive, procedural, and interactional justice. Distributive justice is the perceived fairness of outcomes relative to what others have received (Byrne, 2005; Deutsch, 1985). Distributive justice in the performance feedback setting can be understood as a perception of fairness relative to the work performed (Greenberg, 1986). Procedural justice is concerned with fairness in perceptions of the overall process in which decisions are made (Thibaut & Walker, 1975). Interactional justice, which will not be included in this research

though remains an important factor in justice literature, refers to fairness in treatment and interpersonal communications (Bies & Moag, 1986).

Measurement of fairness perceptions in performance feedback has proved difficult and inconsistent (Keeping et al., 2000), and this research aims to shed light on the role of fairness in this feedback. The three attitudes of satisfaction, acceptance, and fairness have been linked in previous research, but the directionality of these relationships remains unclear (Greenberg & Colquitt, 2013). Context is key in determining if these attitudes are correlates, antecedents, or outcomes to one another (Colquitt et al., 2001). In the context of performance feedback forms, I hypothesize that if feedback format and expectation clarity influence the attitudes of satisfaction and acceptance, it is ultimately due to a change in fairness perceptions.

This rationale is based on procedural justice research, which has shown that people are less likely to feel satisfied or accepting of an outcome if they do not feel it was reached fairly (Colquitt et al., 2001). Fairness has also been shown to explain relationships related to the present study, such as the relationship between perceptions of feedback accuracy and goal setting motivation (Roberson & Stewart, 2006). Therefore, this research connects the attitudes of satisfaction, acceptance, and fairness through mediation.

A variable is considered to be a mediator when it accounts for the relationship between the independent and dependent variable (Baron & Kenney, 1986). In this case, fairness perceptions are the mediating variable. Fairness is expected to mediate the relationship between both independent variables (rating format and expectation clarity) and both dependent variables (satisfaction and acceptance). In order for mediation to occur, a significant relationship between three separate paths in the model must exist. More specifically, Baron et al. (1986) state that mediation occurs when there is a significant relationship between the independent variable and

the mediating variable, between the mediating variable and the dependent variable, and between the independent variable and the dependent variable. The final relationship path, between the independent and the dependent variable, must change from significant to non-significant (or, at least, a reduced effect size) in order for the mediating variable to be considered successful (see Figure 2 and Figure 3).

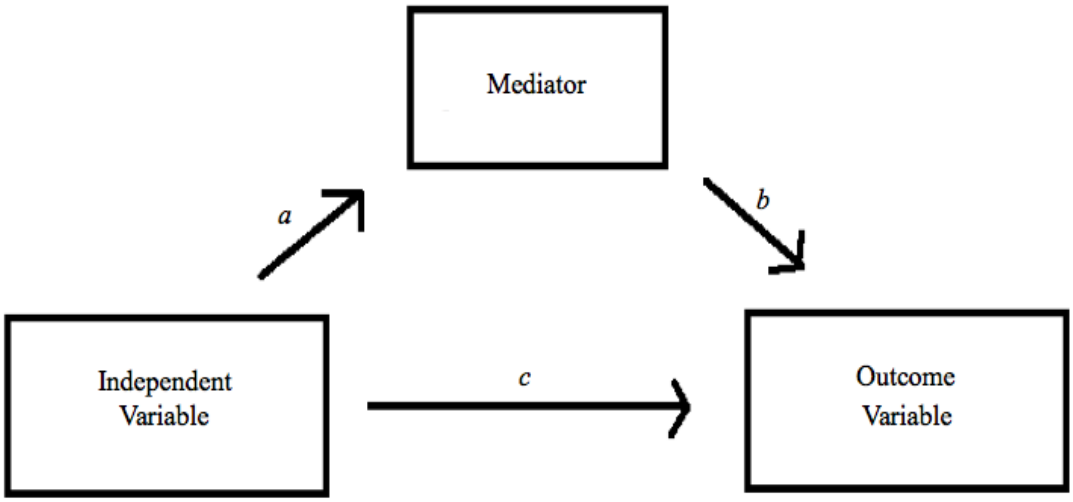


Figure 2. Traditional mediation model.

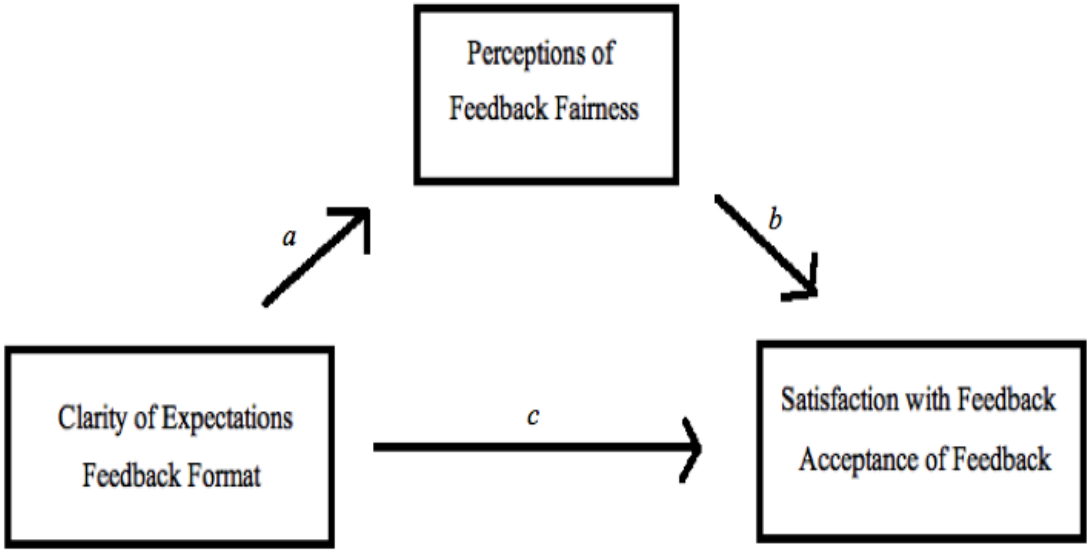


Figure 3. Mediation model with the present study's variables.

H4a: Perceptions of fairness will mediate the relationship between feedback rating format and feedback acceptance.

H4b: Perceptions of fairness will mediate the relationship between feedback rating format and feedback satisfaction.

H4c: Perceptions of fairness will mediate the relationship between clarity of expectations and feedback acceptance.

H4d: Perceptions of fairness will mediate the relationship between clarity of expectations and feedback satisfaction.

METHOD

Participants

Participants were undergraduate students enrolled at a large university in the southwestern United States. They were recruited from five separate undergraduate psychology classes. Individuals who fully completed the survey received extra credit from their instructor. Undergraduate students are the most common type of participants used in policy-capturing (Karren et al., 2002), which is the statistical method used for this research. Further information on policy-capturing will be discussed in detail below.

The survey was housed in Qualtrics and conducted online, with students participating remotely from the computer of their choice. Participants were provided with the survey link and permitted to participate at any time during a two-week period. At the close of the participation window, 280 students had competed the survey. After cleaning the data, there was a final n = 272. 172 (63%) of participants were female, and 258 (95%) of participants were between the ages of 16-24. The majority of participants (79%) identified as White/Caucasian, with 5% identifying as African American, 4% identifying as Asian, and 3% identifying as Latino.

Power Analysis

A power analysis using the statistical software G*Power determined that a sample size of n = 84 would have sufficient power (.80) to test the repeated-measures MANOVA, and a sample size of n = 168 would provide enough power to detect a two-way interaction. This was based on an expected effect size of $f^2 = .1$, $p < .05$, with four groups and two instances of repeated measures. The expected correlation of responses among the repeated measures was approximately .8; as both measures will consist of participants rating nearly identical sets of

forms, their attitudinal responses should be nearly identical as well. More detail on this accuracy of this assumption will be provided in the results section. The estimated small-to-medium effect size was determined based on Cohen's (1988) standards for small (.02), medium (.15), and large (.35) effect sizes for linear models. Additionally, the power value of .80 is taken from Cohen's default recommendation (Cohen, 1992).

Materials

Policy-Capturing. Policy-capturing is a recognized method for determining the influence of specific characteristics on an organizational process (Aiman-Smith, Scullen, & Barr, 2002). This technique has been used in a wide range of organizational research, including compensation (Sherer, Schwab, & Heneman, 1987), job analysis (Sanchez & Levine, 1989), and motivation (Karren et al., 2002; Zedeck; 1977). It allows researchers to capture individual evaluative judgments, and determine the importance of various cues (Karren et al., 2002). It is therefore a useful method to identify the characteristics that influence a ratee's preference for feedback type.

*Explanation of policy-capturing***.** Policy-capturing consists of presenting a series of scenarios to participants, with specific cues in each scenario that represent each predictor variable (Aiman-Smith et al., 2002). In this research context, there are four possible scenarios that represent the two value levels of each cue: high/low clarity of expectations, and numerical/text rating format. The decision outcomes are then regressed on the cues. In this case, the decision outcomes are the participant's attitudes of fairness, satisfaction, and acceptance towards a specific feedback form. Policy-capturing provides insight into which feedback form is most preferred by participants, out of the four possible permutations. As is recommended by best practices, extant literature and popular press were examined as a guide for what the general

scenario of a feedback form should look like (Aiman-Smith et al., 2002). Feedback was also incorporated from subject matter experts.

Policy-capturing can be used to identify the decision-making outcomes of an individual person, also known as idiographic research, or aggregated tendencies across multiple people, known as nomothetic research. This distinction allows researchers to interpret findings more accurately, as well as select the best statistical technique to analyze data (Aiman-Smith et al., 2002). In sync with the majority of policy-capturing research, the present study asks a nomothetic question with the goal of identifying general tendencies across many participants (for more examples of nomothetic policy-capturing research, see Hitt & Barr, 1989; Thoms, McMasters, Roberts, & Dombkowski, 1999).

In order for policy-capturing processes to be meaningful, the scenarios and cues should be as realistic and representative to the participants as possible (Aiman-Smith et al., 2002; Karren et al., 2002). The issue of realism is not a significant obstacle in this context, as participants are being asked to provide opinions about the feedback form, not the content of the feedback. However, realism would potentially be an issue if participants were being asked to imagine that the feedback content was reflective of their own performance, or if the feedback was presented in an atypical manner (Hair, Anderson, Tatham, & Black, 1992). Examples of misrepresentative feedback forms are those containing highly inappropriate language, unintelligible information, or highly irrelevant dimensions.

Policy-capturing results can be questionable if the participants are being asked to weigh in on unfamiliar topics (Aiman-Smith et al., 2002). Researchers cannot confidently use results from a naïve subject pool to infer understanding of a more experienced group. An example of this issue is using student participants to determine what seasoned hiring managers consider

when evaluating job applicants. Therefore, in order to maximize representativeness, the topic of the policy-capturing vignettes in this study needed to be familiar to the participant pool (college students) while still being related to the larger topic at hand (performance feedback forms). To meet these two requirements, the vignettes are feedback forms that will potentially be used by university instructors to provide student feedback. The rating format and design are similar to the evaluations that students are required to fill out at the end of each course, which should heighten the buy-in and familiarity of the form and task. Participants were also informed that the feedback forms have the potential to be used in future courses, which should increase buy-in.

*Execution of policy-capturing***.** The most commonly used medium in policy-capturing is written scenarios, though any rating format that is realistic and effective may be used (Aiman-Smith et al., 2002). As most formal performance feedback is written for documentation purposes, this study presented the feedback forms in written form online. The feedback evaluations that participants complete for their actual university courses have recently moved to an online platform, thus making a written online rating format the most familiar and realistic option.

Most experts recommend that administration of this research type should not exceed one hour, due to possible fatigue and/or loss of interest (Aiman-Smith et al., 2002; Cooksey, 1996). Based on pilot study results, survey completion was estimated to take a maximum of 45 minutes and an average of 25 minutes. This process was therefore within the recommended administration range, and participants should not have experienced the feelings of boredom or exhaustion that are common drawbacks to fully crossed designs in policy-capturing (Graham & Cable, 2001).

As policy-capturing often requires participants to make certain assumptions about the scenarios they are presented with, it is important to inform and prepare subjects about the task at

hand. A set of informative instructions enables the participants to interpret the scenarios from the researcher's intended perspective. It also promotes buy-in from participants, as it provides a rich background and encourages the subjects to treat the scenarios as real life (Aiman-Smith et al., 2002). Therefore, the content and design of the feedback form will be discussed in detail next.

Development of performance feedback form. The feedback forms and instructions presented to participants were created specifically for this study (see Appendix A-I). The design layout was based on a published and publically available performance feedback form used in the United States Air Force (see Appendix J), modified with input from subject matter experts (n = 10). The content of performance dimensions was drawn from general course expectations in undergraduate syllabi to maximize realism.

There were four variations of forms to test each cue. In order to increase power within policy-capturing studies, scenarios are often repeated throughout a given experiment. This allows the researcher to obtain multiple data points for the same variable (Graham & Cable, 2001). Therefore, the four variations of forms were duplicated to create a second set of separate but equal feedback forms. Specifically, the first set of four feedback forms differed only by using small squares corresponding to the rating format options (see Appendix B, D, F, H). The second set of feedback forms used small circles in place of the squares (see Appendix C, E, G, I). This minor difference was assumed to be just noticeable enough to reduce monotony of rating the same form twice, yet not meaningfully different enough to result in response variance. This assumption was tested within the pilot study, and will be discussed in more detail below.

The creation of a second set of four forms resulted in a total of eight feedback forms. These eight forms were pilot-tested to a group of subject-matter experts (n = 10) to ensure the quality of each manipulation, as well as the appropriateness of the form's design and layout.

Pilot test results indicated that the independent variables were being represented appropriately, and the nearly identical means between the two sets of forms indicated that the two sets could be considered different but equal (see Table 1).

Table 1. Means and Standard Deviations of Pilot Responses by SME's Given Two Versions of Four Feedback Types (with Standard Deviations in Parentheses).

		DV	DV	Mediator
Form Type	*n*	Satisfactory text rating format	Satisfactory numerical rating format	Satisfactory high clarity
High Clarity and Text (V1)	9	3.89 (.83)	1.67 (1.00)	3.56 (.73)
High Clarity and Text (V2)	9	3.44 (.73)	1.56 (.73)	3.00 (.71)
High Clarity and Numerical (V1)	9	1.50 (1.07)	4.00 (.00)	3.50 (.76)
High Clarity and Numerical (V2)	9	1.22 (.44)	3.56 (.88)	3.00 (.71)
Low Clarity and Text (V1)	9	3.38 (.92)	1.38 (.74)	2.38 (.52)
Low Clarity and Text (V2)	9	3.33 (1.12)	1.11 (.33)	2.22 (.67)
Low Clarity and Numerical (V1)	9	1.33 (.71)	3.56 (.73)	2.00 (.50)
Low Clarity and Numerical (V2)	9	1.33 (.71)	3.56 (.73)	2.22 (.67)

Note. Response options were a 4-point Likert scale, where 1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree.

In addition to these eight forms, eight distractor forms were included in the survey. Distractors are helpful at preventing demand effects from influencing the participants (Aiman-Smith et al., 2002). With a relatively small amount of cue variables and forms being tested, participants may realize the research hypothesis and then answer in a manner that does not reflect their true attitude (Shimp, Hyatt, & Snyder, 1991). Distractor forms are a simple way to keep the participant from guessing at the nature of the study. These distractor forms were similar

to the real forms, but included modified pieces of information and/or presented the information in a different layout (see Appendix K-R). The responses to these eight distractor forms were not analyzed in the results of the study.

In summary, each participant rated a total of sixteen forms. Two sets of four forms were used to capture each cue variable: text and numerical rating format, and high or low clarity of expectations. Eight additional forms were interspersed to distract the participants from identifying the independent variables of interest. These sixteen forms were randomized within the survey software Qualtrics to prevent any systematic response fatigue from influencing results.

Feedback rating format*:* The independent variable of feedback rating format was represented through two possible conditions: text rating format and numerical rating format. The text rating format condition used the following words to categorize performance levels: "poor", "below average", "average", "above average", and "exceptional". The numerical rating format represented these levels using scale numbers from "1" through "5". To avoid the issue of potentially providing additional information in the numerical condition, each rating format was comprised of five options and only whole numbers were used within the numerical condition. This satisfied the criticisms of previously used numerical rating formats being preferred due of a wider defined range of possible information (i.e. decimal-point options), as seen in the Atwater et al. (2006) study. These manipulations were presented to subject matter experts and found to be satisfactory representations of the rating format variable.

Clarity of feedback standards. Using Cardy et al.'s (2004) interpretation, expectation clarity was divided into three dimensions. The first aspect of expectation clarity is an indication of the relevant performance dimensions. For example, this study will focus on three dimensions

of performance: tests, quizzes, and projects. The second and third aspects of clarity are closely tied: explaining the possible range of performance levels, and providing examples of behaviors associated with those different levels. For example, on a performance dimension of "attendance", clear explanation of performance levels would be "rarely attends class", "usually attends class", and "always attends class". Examples of behaviors associated with the lowest level of performance would be "attends class less than once a week".

The high clarity measure provides detailed information for all three of these expectation dimensions. As it did not seem realistic for the low clarity condition to be completely blank or unintelligible, the low clarity measure consists of the performance dimension names and a vague reference to the range of performance levels. For example, a high clarity example of low attendance is "attending class less than once per week on average". A low clarity example of the same behavior is "rarely attending class". These manipulations were presented to subject matter experts and found to be satisfactory representations of the expectation clarity variable (see Table 1).

Satisfaction. The dependent variable of feedback satisfaction was assessed using the existing measures by Thurston & McNall (2009), with a reliability of .90. The reliability for the scale in this present study was $\alpha = .75$. The items were modified slightly to reflect performance feedback rather than performance appraisal systems. The three items are as follows: "I would be satisfied with receiving this performance feedback form", "I would be satisfied if my performance was evaluated using this feedback form", and "I am satisfied with the way this form presents performance feedback." Subjects responded to a 5-point Likert scale, with 1 = highly disagree to 5 = highly agree (see Appendix S for all scale items).

Acceptance. The dependent variable of feedback acceptance was assessed based on the existing measure of feedback accuracy by Stone & Stone (1985), with a reliability of .95. The reliability for the scale in this present study was α = .80. The three items are as follows: "This form could provide an accurate evaluation of class performance", "This form could reflect feedback that is based on true class performance", and "This form is similar to my own idea of class performance feedback". Subjects responded to a 5-point Likert scale, with 1 = highly disagree to 5 = highly agree.

Fairness. The mediating variable of fairness perceptions was measured using Colquitt's (2001) procedural justice scale (α = .85). The reliability for the scale in this present study was α = .79. The scale has been slightly modified to reflect general performance feedback rather than a specific performance appraisal. The three items are as follows: "This form could provide feedback that is free of bias", "This form could provide performance feedback that is based on accurate information", and "This form displays feedback in a consistent manner". Subjects responded to a 5-point Likert scale, with 1 = highly disagree to 5 = highly agree.

Design. I employed a full factorial design, in which all variables were crossed (see Figure 4). This allowed for the analysis of the independent effects of each variable on each participant's decision (Graham & Cable, 2001). As the number of manipulated variables was small (two variables at two levels each), the common issue in policy-capturing of exhausting participants was not a concern. A fully crossed design only yielded four separate forms for participants to rate. Ultimately, eight forms were used (two sets of four forms), including an additional eight distractor forms. Each participant rated every form, making this a fully within-subjects design.

		Rating format	
		Text	*Numerical*
Clarity	*High*	Text/High Clarity	Numerical/High Clarity
	Low	Text/Low Clarity	Numerical/Low Clarity

Figure 4. Factorial design.

In order to achieve a richer range of data, the outcome variables of acceptance and satisfaction were treated as numeric (using a Likert scale to indicate degree of acceptance and satisfaction) rather than categorical (e.g. indicating "yes" or "no" for satisfaction and acceptance).

Procedure

Each participant took the survey online and rated all sixteen feedback forms. In order to increase psychological buy-in, participants were informed that their opinions might be incorporated into feedback forms used in their future courses. The instructions for the scenarios in this study were as follows:

"The Psychology Department at Colorado State University is developing a form for instructors to provide students with feedback on individual class performance. This form will potentially be used by all programs and colleges within the university. We are asking for student input on these feedback forms. Your opinion will directly contribute to the final product.

You will be shown 16 samples of feedback forms, and asked to provide your reaction to these forms. Please be able to explain which aspects of each form are impacting your reactions."

Participants viewed one form at a time in a randomized order. Each form was followed by ten multiple-choice questions to measure acceptance, satisfaction and the mediating variable

of fairness. After rating all sixteen forms, participants were asked to rank-order their preferred forms out of four options. The survey ended with demographic questions and a brief thank-you message.

Data Cleaning. This study required participants to thoughtfully reflect on several feedback forms and indicate their attitudinal reactions and preferences. To reduce measurement error that could occur due to invariant responding, the recommendations of Huang, Curran, Keeney, Poposki, & DeShon (2012) were followed. It was unlikely that participants would have the exact same responses to both distractor forms and real forms, as the distractor forms were generally varied and uniquely formatted. Eight participants (3%) were eliminated due to response invariance; specifically, participants who selected the same Likert scale response for at least twelve of the sixteen forms (75%). Their data was removed prior to conducting any analyses.

P-value cutoff. For the purpose of this research, a *p*-value of .1 or lower is considered significant. As the focus of this study's results will be on effect sizes rather than significance levels, a more liberal cutoff allows for the examination of results that would traditionally not meet the criteria for further investigation. Though a more liberal cutoff can increase the risk of committing a Type 1 error, there is a history of support for researchers adjusting the cutoff value from .05 to .10, particularly for identifying trends within the social sciences (Kline, 2004; Krivoshey, 1975; Larson-Hall, 2010; Murphy & Myors, 2004).

RESULTS

Descriptive Statistics

As discussed in the Methods section, the eight distractor forms were removed prior to data analysis. This resulted in data for two sets of four forms. A correlation matrix of the variables, as well as the means and standard deviations of responses to each form, were examined first to identify data trends (see Tables 2 and 3, and Appendix T and U for the full version inclusion of both form sets). While the significance of these trends will be explored in detail by the multivariate and univariate analyses to follow, these basic results can inform our expectations of the relationships prior to running more sophisticated analyses. Before interpreting the descriptive statistics, it is important to note the narrow range of mean responses. While most of the mean differences may appear to be slight, they are negatively skewed on a five-point scale. These restrictions lend support to interpreting the slight differences with careful consideration.

Table 2. Correlation Matrix of all Dependent and Mediating Variables.

	Text/LC	Text/HC	Num/LC	Num/HC	Satisfaction	Acceptance	Fairness
Text/LC	____						
Text/HC	.27*	____					
Num/LC	.50*	.14*	____				
Num/HC	.21*	.64*	.32*	____			
Satisfaction	.60*	.69*	.62*	.72*	____		
Acceptance	.64*	.70*	.62*	.73*	.87*	____	
Fairness	.57*	.61*	.53*	.63*	.58*	.65*	____

**$p < .001$.

Table 3. Mean of Responses Regarding Feedback Forms on Satisfaction, Acceptance, and Fairness Scales.

Form Type	*n*	DV (SD) Satisfaction	DV (SD) Acceptance	Mediator (SD) Fairness
High Clarity and Text	272	3.42 (.94)	3.39 (.87)	3.68 (.79)
High Clarity and Numerical	272	3.34 (.98)	3.38 (.88)	3.69 (.77)
Low Clarity and Text	272	2.78 (1.00)	2.86 (.96)	3.27 (.86)
Low Clarity and Numerical	272	2.72 (.98)	2.83 (.93)	3.24 (.87)

**Note: Maximum response possible was 5.00 (Strongly Agree)*

Correlation Matrix. The matrix of correlations between the variables reveals that all variables are positively correlated with each other. The strong positive correlations between satisfaction and acceptance were expected, as these measures are often combined within feedback research to represent general appraisal reactions. However, the extent of this positive correlation (.87) indicates that participants may have interpreted two variables as the same construct, thus making the two redundant. The implications of this will be discussed in further detail within the discussion section. The weak-to-moderate relationships between fairness and the outcome variables raise some initial concern over the likelihood of mediation, as we would expect to see fairness more strongly correlated with the dependent variables in a full mediation model.

Satisfaction Means. Forms that were high in clarity and used a text-based rating format resulted in the highest mean responses of satisfaction. Regardless of the rating format type, participants reported less satisfaction in low clarity conditions ($\bar{x} = 2.75$) than in high clarity conditions ($\bar{x} = 3.38$) (see Figure 9). The mean satisfaction responses were lowest when a feedback form was low in clarity and numerically formatted ($x = 2.71$) (see Figure 10). See Appendix V-W for the mean satisfaction ratings of both sets of forms.

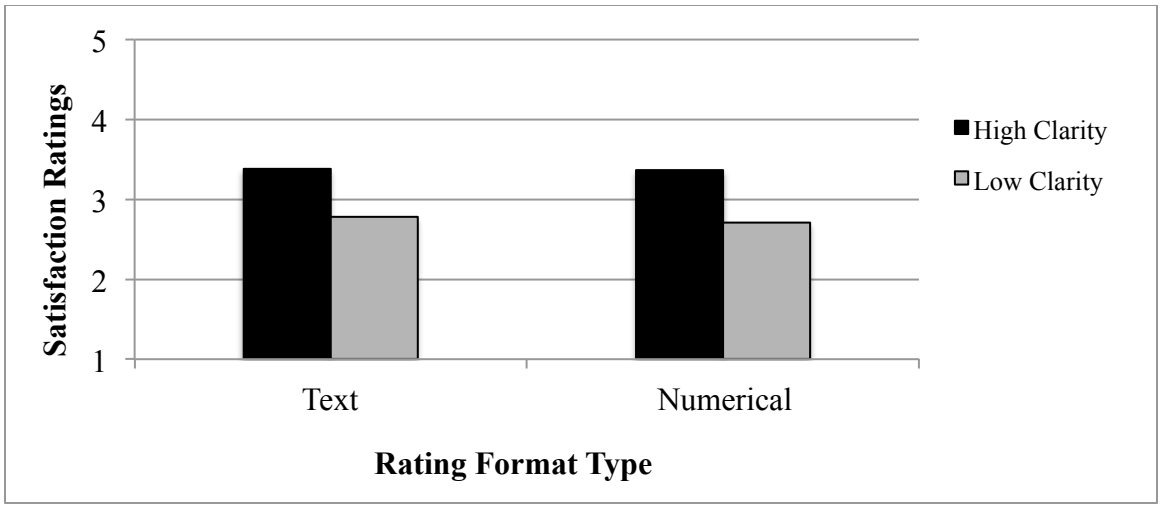


Figure 9. Mean Satisfaction Ratings By Clarity Type Across Rating Format Type.

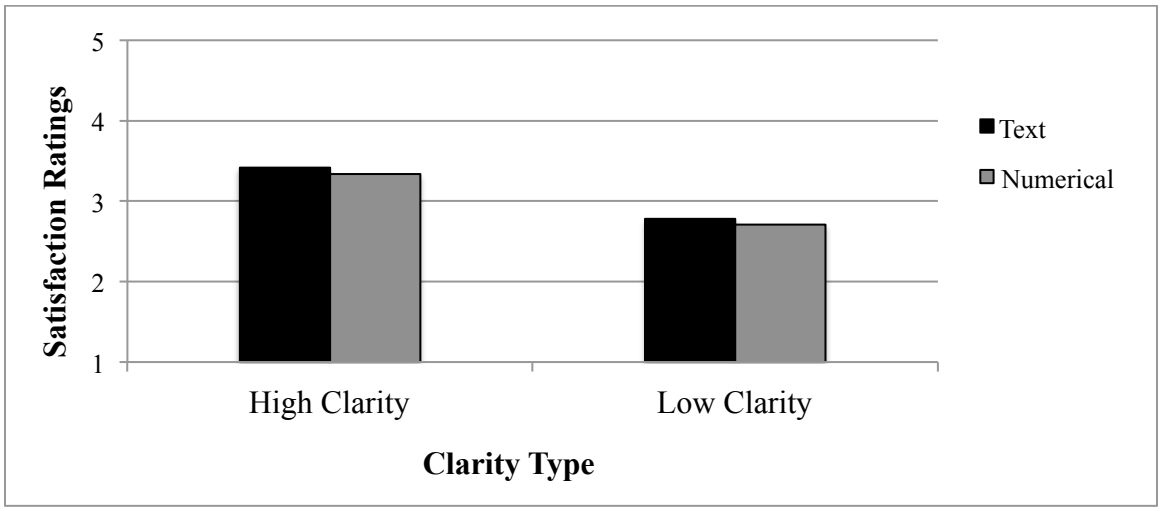


Figure 10. Mean Satisfaction Ratings By Rating Format Type Across Clarity Type.

Acceptance Means. On average, the means of acceptance were slightly higher ($\bar{x}$ = 3.11) than were the means for satisfaction ($\bar{x}$ = 3.06). While this difference could be due to chance, it is possible that participants were more easily accepting of the feedback forms in general, which provides support for measuring acceptance and satisfaction as separate variables. Yet, the mean responses for acceptance follow trends similar to the satisfaction responses.

Participants reported higher acceptance of the feedback forms in high clarity conditions ($\bar{x}$ = 3.38) than low clarity conditions ($\bar{x}$ = 2.83), regardless of the rating format type. Similar to the satisfaction responses, participant acceptance of the feedback forms appeared to be most strongly influenced by the form's clarity (see Figure 11). The mean acceptance responses were lowest when a feedback form was low in clarity and numerically formatted (x = 2.81). These mean differences were even smaller than the differences found within satisfaction responses, differing by only .03 units, but still provide evidence of a consistent response pattern across attitudes by rating format (see Figure 12). See Appendix X-Y for the mean acceptance ratings of both sets of forms.

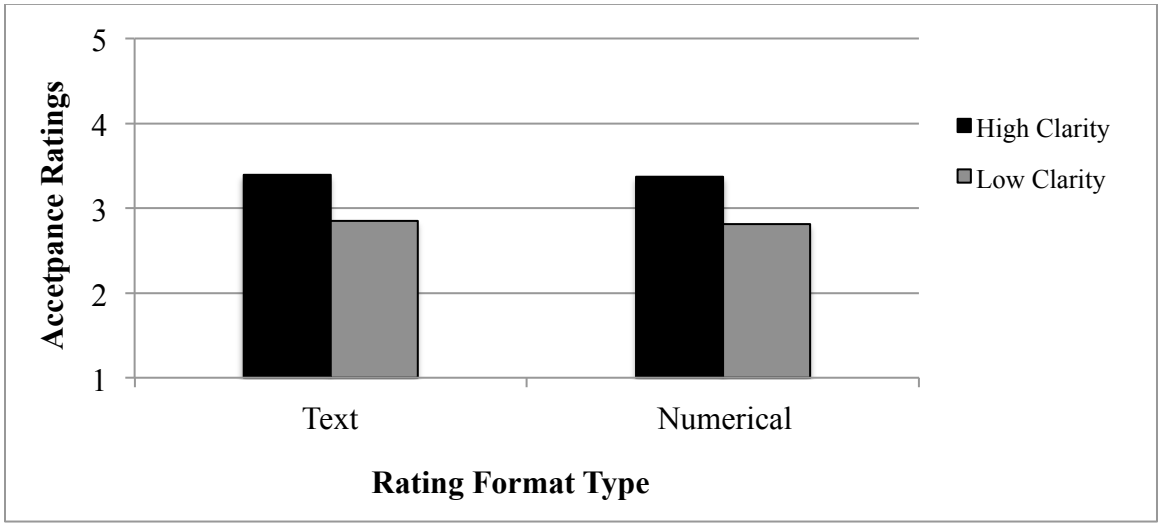


Figure 11. Mean Acceptance Ratings By Clarity Type Across Rating Format Type.

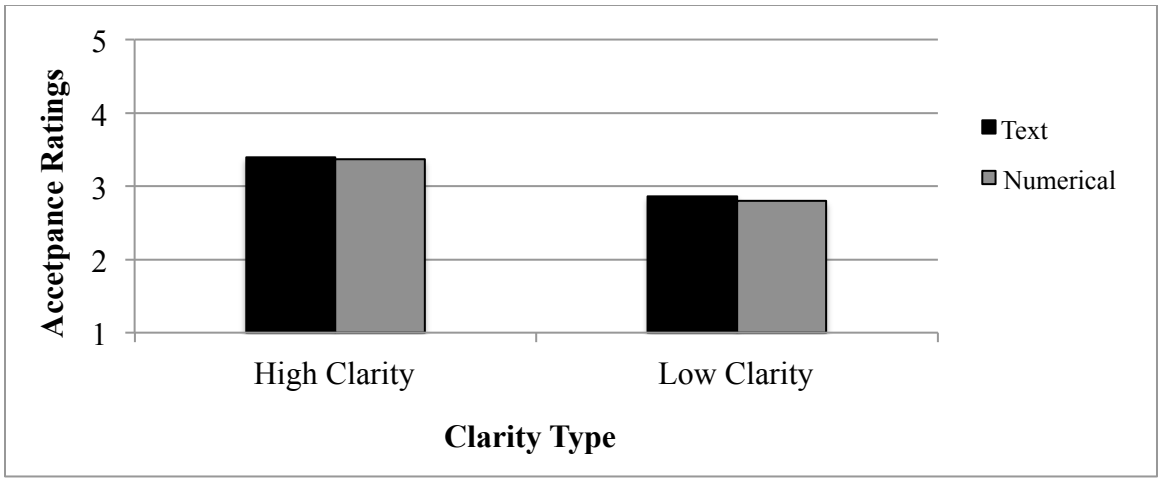


Figure 12. Mean Acceptance Ratings By Rating Format Type Across Clarity Type.

Implications for Both Attitudes. Based on the descriptive statistics, both independent variables of rating format and clarity of expectations influenced participant's satisfaction and acceptance of the feedback forms. Clarity appeared to affect the means more than rating format, but the mean responses for both attitudes follow the same pattern across clarity and rating format.

Fairness Means. Though it is not an outcome variable, the mean and standard deviations of the fairness variable can provide information about the likelihood of mediation. As will be explained in detail during the analyses, the relationship between the independent variables and the mediator is an important first step in identifying mediation.

The overall fairness means are noticeably higher ($\bar{x} = 3.47$) and contain smaller standard deviations than the satisfaction or acceptance responses (see Figures 15 and 16, and Appendix AB-AC for the overall mean outcome ratings of both sets of forms). It is unclear why participants would have an easier time endorsing all feedback forms as fair, though potential explanations will be explored in the Discussion section. The clarity conditions map on to previous results, with the low clarity and numerical rating format condition receiving the lowest

fairness means (x = 2.24). Notably, though high clarity forms were still associated with higher means ($\bar{x}$ = 3.68) than low clarity forms ($\bar{x}$ = 3.25), the difference between these two clarity measures was considerably less than in the satisfaction and acceptance clarity differences (see Figure 13). In general, there was a more restricted range of mean responses to the fairness questions. Mean fairness attitudes were not noticeably different depending on form type, which makes fairness an unlikely (though not impossible) mediator (see Figure 14). See Appendix Z-AA for the mean fairness ratings of both sets of forms.

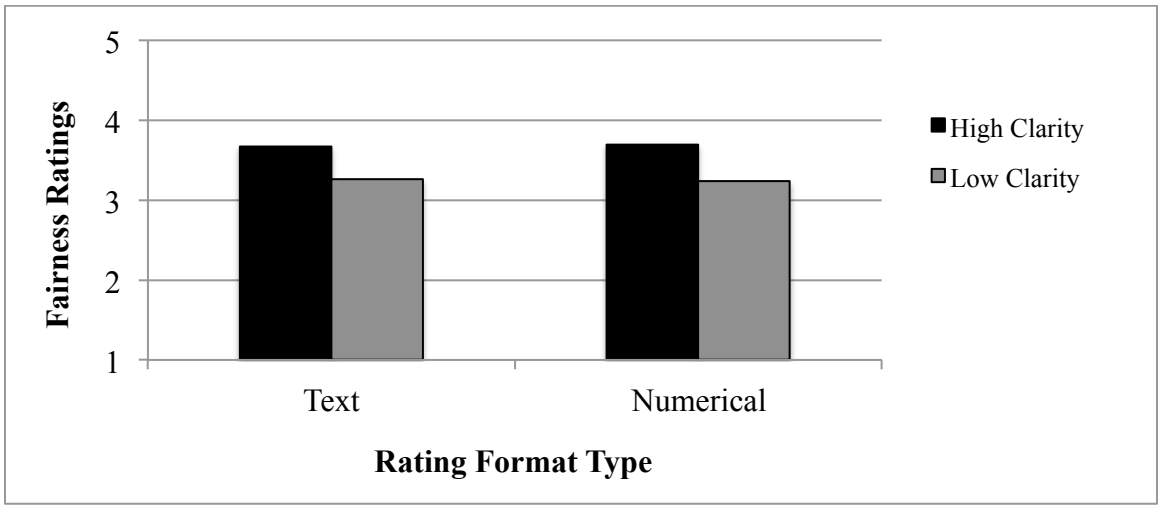


Figure 13. Mean Fairness Ratings by Clarity across Rating Format Type.

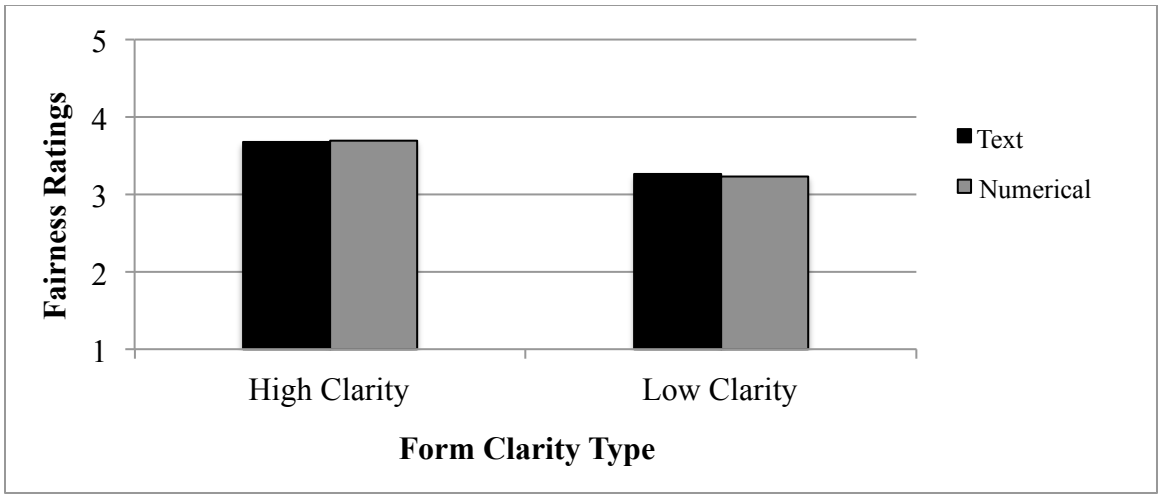


Figure 14. Mean Fairness Ratings by Rating Format across Clarity Type.

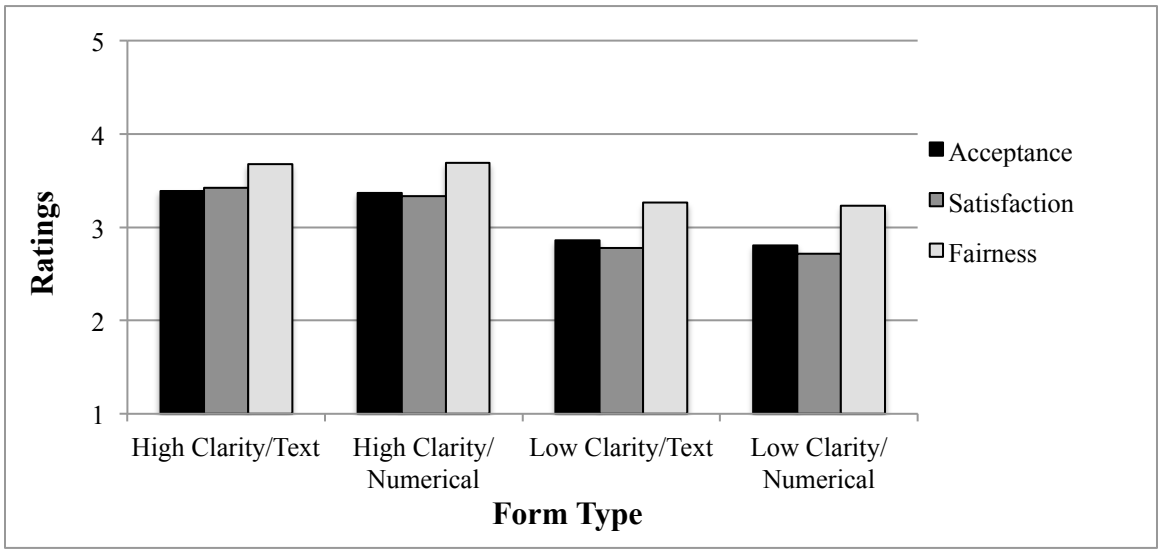


Figure 15. Mean Satisfaction, Acceptance, and Fairness Ratings Across Feedback Form Type.

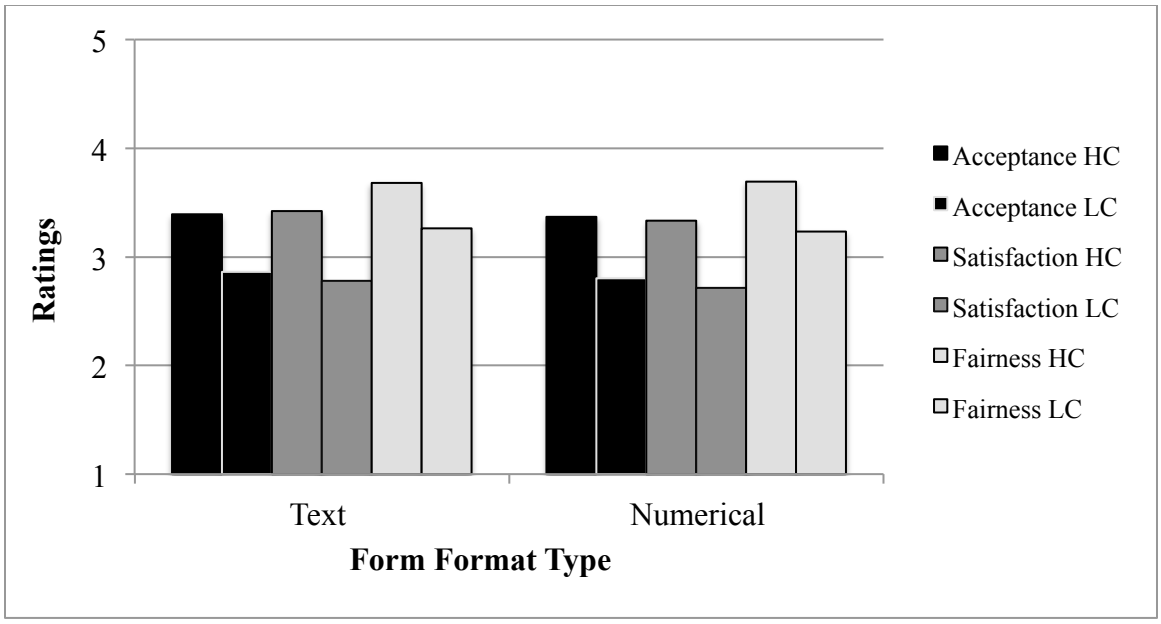


Figure 16. Mean Satisfaction, Acceptance, and Fairness Ratings By Clarity Across Rating Format Type.

Assumption Checks

Outliers. A test for outliers in all three attitudinal measures (satisfaction, acceptance, and fairness) revealed no responses outside of three standard deviations from the mean, which was the *a priori* threshold for elimination given the restricted range of responses.

Sphericity and Normality. Results for Mauchley's Test of Sphericity indicated no violations of sphericity assumptions. Tests of normality for all three attitudinal responses show statistical significance ($p < .05$), indicating that the assumption of normality has been violated. While this test statistic can be useful in deciding whether or not to move forward with analyses, both univariate and multivariate tests of normality are highly sensitive to departures from a normal distribution and the F test is known for its robustness to violations of normality (Hancock et al., 2010). Therefore, despite the significance indications of normality violations, I acknowledged this limitation and moved forward with the analysis.

Paired sample t-tests. When the second set of four feedback forms were developed, they were created to be identical to the first set, barring a small graphical difference. Specifically, the

first set of four feedback forms had small circles corresponding to the rating format options (see Figure 17 for an example, and Appendix B, D, F, H for the full version). The second set of feedback forms had small squares in place of the circles (see Figure 18 for an example, and Appendix C, E, G, I for the full version). This minor difference was created to be just noticeable enough to reduce monotony of rating the same form twice, yet not meaningfully different enough to result in response variance.

Figure 17. Example of feedback form in Set 1 (square format).

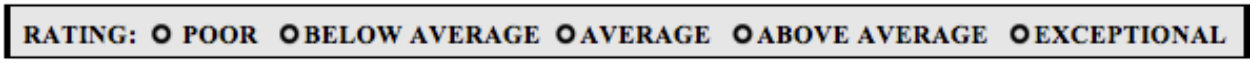


Figure 18. Example of feedback form in Set 2 (circle format).

To investigate the accuracy of the assumption that participants would rate these two sets of forms identically, I conducted three sets of paired t-tests for satisfaction, acceptance, and fairness. Paired t-tests are useful tools for identifying significant response differences between repeated measures. The four forms in Set 1 were paired with their Set 2 counterparts to investigate the assumption that participants would respond similarly to these near-identical forms.

Across the dependent and mediating variables, the two sets of four forms did not yield significantly different responses within subjects ($p > .05$), with t-values ranging from .041 to 1.39. Each pair of form responses had an average correlation of .75 ($p < .001$). The item-total mean differences between each pair ranged from .002-.100 units, with an average mean difference between pairs of .038. This means that, on average, the repeated forms within Set 2

differed from their Set 1 counterpart by less than .04 units. This supported the interpretation that the responses in the first set of forms were not meaningfully different from the responses in the second set of forms, and that participants viewed the two sets as similar despite the small graphical differences. Given this interpretation, the two sets of four forms were combined and discussed as one set of four forms for the remainder of the analyses and discussion.

Overview of Analyses

The hypotheses for this data were analyzed using a repeated-measures mulitvariate analysis of variance (MANOVA). This technique was most appropriate given two independent variables with two levels each, and more than one dependent variable being measured at two time points (Green & Salkind, 2003; Hancock & Mueller, 2010). A multivariate analysis of variance was more appropriate than repeated univariate tests, as MANOVA reduces the likelihood of Type 1 error. MANOVA tests also account for a potential statistically significant combination of independent variables, even in cases when the individual ANOVA tests would not show statistical significance. This provides information regarding the effect of the predictor variables as a whole, rather than simply their effect in isolation of each other. Repeated-measures designs tend to result in a higher degree of parameter estimation precision and inferential power in situations of smaller-than-preferred sample sizes (Hancock et al., 2010). Though sample size was ultimately not of concern for this study, this type of design remains an appropriate fit.

The repeated-measures MANOVA included three independent variables, which consisted of rating format (at two levels), clarity (at two levels), and repetition (at two levels). Repetition was included to account for the second set of highly similar forms, but this variable was not examined during the data analysis stage. The results of the MANOVA reveal if

participant reactions to the forms vary across the levels of each factor. If MANOVA results are statistically significant, then it is permissible to examine the univariate tests to obtain a more detailed understanding of main effects. Therefore, the multivariate test results will be reported before potentially progressing to the univariate test results.

Results of Analyses

Hypothesis 1 stated that participants would prefer feedback forms with numerical ratings to forms with text ratings. The results of the repeated-measures MANOVA test showed a non-significant effect for rating format, Wilks' $\lambda = .991$, $F(3, 269) = .849$, $p = .468$, $\eta^2 = .01$. Therefore, Hypothesis 1 was not supported.

Hypothesis 2 stated that participants would prefer feedback forms that were high in clarity of expectations to forms with low clarity of expectations. The results of the repeated-measures MANOVA test showed a significant multivariate main effect for clarity, Wilks' $\lambda = .977$, $F(3, 269) = 2.13$, $p = .097$, $\eta^2 = .02$. Therefore, based on the marginally greater effect size and statistically significant multivariate test results, the variable of clarity warrants further examination on a univariate test level to determine if its effect differs by each attitudinal outcome.

A significant but small univariate main effect for clarity was found for satisfaction, $F(1, 271) = 3.94$, $p = .05$, $\eta = .01$. Univariate tests did not reveal a significant relationship between clarity and acceptance, $F(1, 271) = 1.27$, $p = .27$, $\eta = .01$. While the means did reveal differences in both attitudinal responses by clarity condition and the effect sizes for both main effects are similar, only satisfaction attained statistical significance. As both attitudes of satisfaction and acceptance were expected to be statistically significant, Hypothesis 2 was only partially supported.

Hypothesis 3 stated that there would be an interaction between rating format and expectation clarity. It was expected that participants would prefer numerically based feedback forms in low clarity conditions, but that preference would lessen in high clarity conditions. The results of the repeated-measures MANOVA test showed a non-significant multivariate effect for the interaction between rating format and clarity, Wilks' $\lambda = .994$, $F(3, 269) = .578$, $p = .630$, $\eta^2 = .01$. Therefore, Hypothesis 3 was not supported.

The final set of hypotheses, *4a-4d*, stated that perceptions of fairness would mediate the relationship between both of the independent variables (feedback rating format and clarity of expectations) and both of the dependent variables (feedback form acceptance and feedback form satisfaction). These hypotheses could be tested in a number of ways, such as structural equation modeling or hierarchical linear modeling. The present analysis used Baron & Kenny's (1986) recommendations, which are widely cited and commonly used to determine mediation. Under their recommendation, evidence of mediation therefore relies on a statistically significant relationship between the independent variables and the mediator, as well as the mediator and the dependent variable (see Figure 2 and Figure 3). In addition to the significant relationships of path *a* and path *b*, a previously significant relationship between the independent and dependent variable(s), illustrated as path *c*, must be reduced (for partial mediation) or eliminated (for full mediation) as a result of the aforementioned relationships (path *a* and path *b*).

To test these separate paths, a series of regression models were estimated. Following best practices, the first test was a regression of the mediator (fairness) on the independent variables of clarity and rating format (Baron et al., 1986). This relationship was non-significant for both clarity, $F(1, 271) = .07$, $p = .79$, and rating format, $F(1, 271) = 1.42$, $p = .24$.

If either of these main effects showed statistical significance, mediation would have been tested further by regressing the dependent variables on the independent variables of rating format and clarity, and lastly, regressing the dependent variables on rating format, clarity, and the mediator of fairness. However, as the first regression model yielded non-significant results, mediation should not be investigated further and cannot be inferred (Baron et al., 1986). Without a statistically significant relationship between the independent variable and the mediator, the proposed model of mediation is not possible and Hypotheses 4a-4d were not supported. It is important to note that while power could have potentially been an issue during the later stages of this mediation investigation, the current sample of $n = 272$ was sufficient for detecting the bivariate relationship (path *a*) of the mediation. Therefore, there was no basis to support moving forward with the proposed mediation given the non-significant results of the first regression model.

Post-hoc Analyses on Participant Reactions to Forms

Open-ended responses. In addition to measuring attitudinal reactions with Likert scale items, participants were asked to provide an explanation of their overall ratings for each form. This explanatory option provided participants with the ability to elaborate on their rating decisions. Participants were given a list of seven options from which to choose, and were instructed to select as many as they found applicable. These options included both positive and negative interpretations of the forms, and mapped on to the independent variables as well as the mediating variable of fairness. The seven options were as follows: "The rating options are clear", "The explanations of high, average and low performance are clear", "The meaning of the rating options was clearly explained", "I would trust instructors to use this form appropriately", "The rating options are not clear enough", The explanation of high, average, and low performance are

unclear", "The form is too busy with information", and "This form seems unfair." Figures 16 and 17 represent the range of selected options by form. For clarity of interpretation, the favorable responses and unfavorable responses were graphed separately.

Shown in Figure 19, high clarity forms had the highest frequency of positive reactions. 65% of the "favorable" response selections were in reference to high clarity forms, indicating that participants viewed these forms as either high in clarity, the preferred rating format, or high in fairness. Fewer positive options were selected in the low clarity conditions, with only 35% of the total favorable responses selections being chosen to describe low clarity conditions. These open responses support the interpretation that level of clarity had a meaningful impact on participant reactions to forms, and could be considered supporting evidence for Hypothesis 2.

It is interesting to note the response trends for the rating format option. Based on Hypothesis 1, the frequency in which participated selected the "clear rating format" option was expected to differ based only on the rating format condition. However, this option was most frequently chosen in high clarity conditions, regardless of the rating format condition. It is possible that the high clarity condition induced a halo effect, which made aspects of the form that were unrelated to clarity seem more preferable as well.

Shown in Figure 20, the unfavorable reactions to forms have an overall similar pattern. 67% of the "unfavorable" response selections came from participants in the low clarity conditions. Participants in these conditions described their forms as unfairness, low in clarity, and containing poor rating formats. Note that the second low clarity response option appeared difficult to endorse. This is likely due to a lack of relevancy of the response option ("This form is too busy with information"). This option was not as clearly related to the low clarity condition, which likely contributed to its infrequent selection.

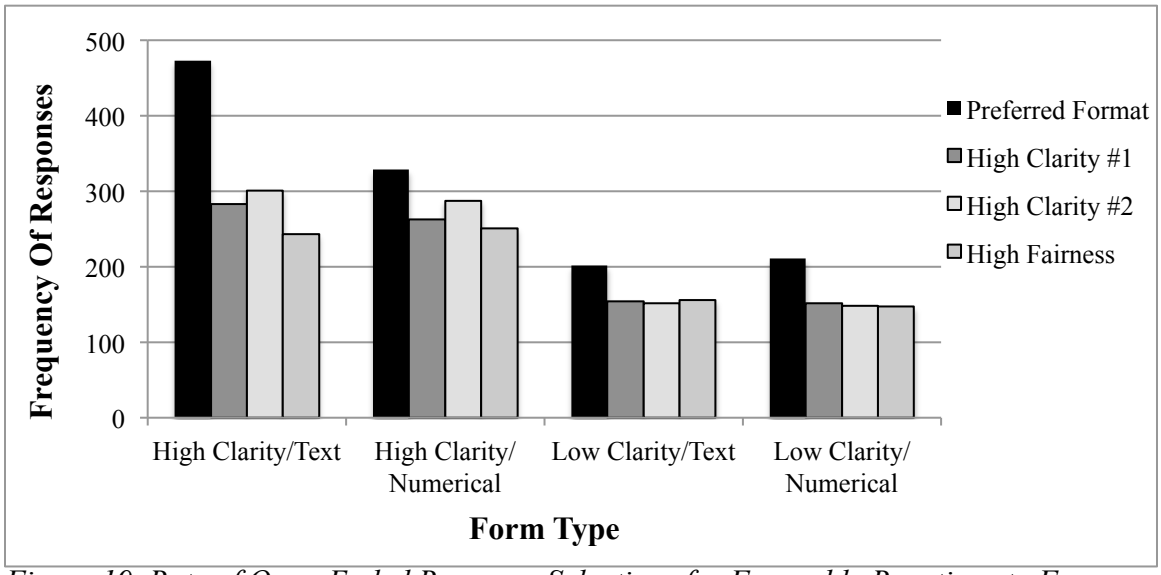


Figure 19. Rate of Open-Ended Response Selections for Favorable Reactions to Forms.

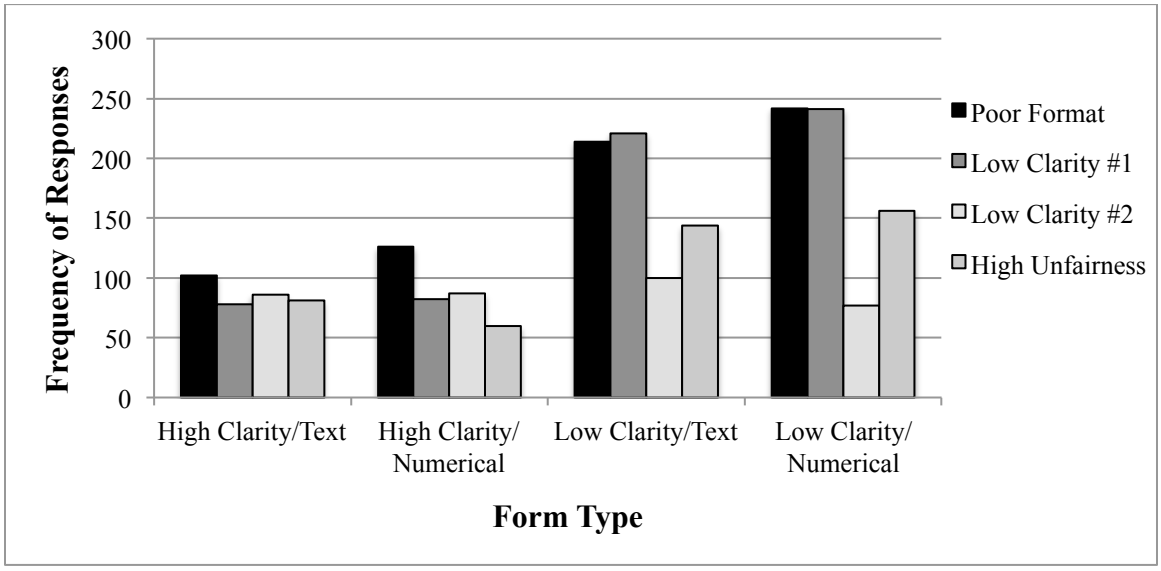


Figure 20. Rate of Open-Ended Response Selections for Unfavorable Reactions to Forms.

Ranking of Preference. At the end of the survey, participants were asked to rank order their preference of the four different forms. They were presented with the four types of forms in a randomized order, and used a drag-and-drop function on their computer to sort the forms in order of preference from first to last. Figure 21 shows that the form rankings follow the trends

established by the earlier statistical analyses. The high clarity and text rating format forms earned 50% of the first-place votes, making them the most preferred type of form. The high clarity and numerical rated forms were ranked as the second-highest preferred form, earning 30% of the second-place votes. The third-place rank was low clarity and text rating forms, with a similar 30% of votes. Lastly, the low clarity and numerically formatted forms earned 51% of last-place votes, putting them as the overwhelmingly least preferred form.

These rankings show an extreme distinction in preference between high clarity/text forms and low clarity/numerical forms, but only a slight distinction in preference between the high clarity/numerical forms and low clarity/text forms. These results imply that participants had a clear idea of their most and least preferred forms, but had a more difficult time ranking their second and third preferences. Despite not showing support for the hypothesized preferred rating format, these results echo the earlier findings that high clarity and text-based rating options are the most preferred feedback form.

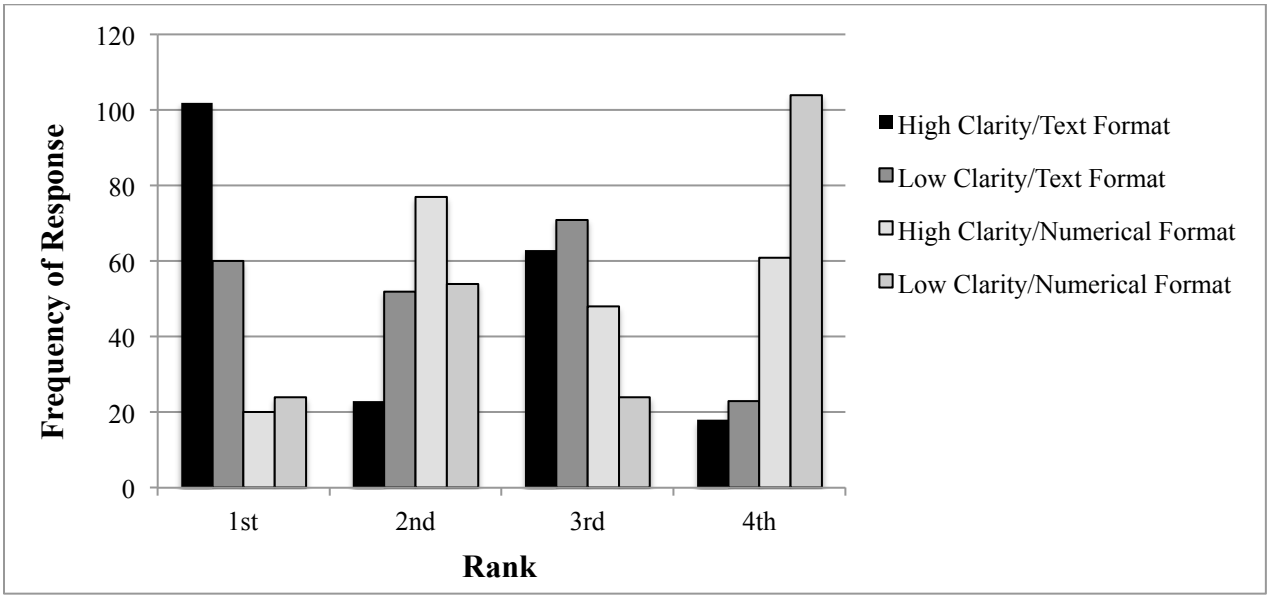


Figure 21. Form Preferences.

Demographics. The collected demographic information on age and gender was analyzed for potential response differences. Participants generally followed the same response patterns independent of age (Appendix AD-AF) or gender (Appendix AG-AI). Any slight differences by the gender demographics were likely due to the relatively small amount of male participants (n = 100), resulting in less stable responses than the females (n = 172). A similar interpretation can be applied for the differences seen in participants over the age of 25 (n = 15).

DISCUSSION

The purpose of the present study was to investigate if certain feedback form features would influence a ratee's preference for the form. Specifically, the study addressed the potential influence of rating format and performance expectation clarity on ratee acceptance and satisfaction, as well as the potential mediating effect of fairness on these relationships. This discussion focuses on each of the hypotheses in turn, as well as limitations and directions for future research. This study had sufficient power to detect the hypothesized group differences for both rating format and expectation clarity, and therefore it is unlikely that the non-significant results are due to a lack of power or insufficient sample size.

Rating Format. It was hypothesized that participants would have more positive reactions to feedback forms that used numerical rating formats than text-based rating formats. Results showed that rating format did not have a significant effect on any attitudes towards feedback forms, $p > .05$. Satisfaction, acceptance, and fairness perceptions of the feedback form were not meaningfully affected by the form's rating format. This is contrary to what was found by Atwater & Brett (2006), whose findings were the basis for this research's hypotheses.

The present findings have several possible explanations. The most likely interpretation of these results is that rating format is not important to ratees when they are forming their attitudes about feedback forms. In the Atwater et al. study, extra information was provided within the numerical format, which may have influenced their results. The present study took precaution to avoid any differences in amount of information between the two rating formats. By carefully standardizing the information between the text and numerical ratings forms, the current study may be a more accurate reflection of the relationship (of lack thereof) between feedback rating

format and form preference. Thus, rather than finding support for the significance of text versus numerical rating formats, this study found the distinction to be unimportant.

This lack of importance is further supported when examining the open-ended response questions. Participants tended to view either rating format option as preferable as long as it as paired with a high clarity condition, and both rating format options were viewed as undesirable when shown in a low clarity context. Therefore, the variable of rating options appeared to be eclipsed by clarity of standards. In both the scale responses and the open-ended responses, rating format did not meaningfully contribute to an understanding of the variance in feedback form preference.

Although this conclusion is contrary to the present study's hypothesis, the statistically non-significant impact of rating format has been supported by past research. While it is generally accepted that a feedback form should include some type of standardized rating system, there may not be much purpose in differentiating between text or numerically-based options. The present study extends the claims made by Landy and Farr (1980), when they concluded that format differences in rating scales do not meaningfully contribute to variance in ratings. Similarly, while ratees may have noticed the differences in rating format options within this study, those observations did not affect their attitudes of satisfaction or acceptance of the feedback form.

Clarity of Expectations. It was hypothesized that participants would prefer feedback forms that had a high clarity of expectations rather than a low clarity of expectations. Though the mean differences indicated a trend that supported this hypothesis and there was a marginally significant multivariate effect ($p = .10$), univariate test results showed only participant satisfaction was influenced by clarity, $p < .05$. This provides partial support that satisfaction and acceptance are distinct variables, as well as partial support for the hypothesis that both

satisfaction and acceptance are affected by a form's clarity of expectations. However, due to the small effect sizes, it is advisable to interpret the importance of these findings with caution. Furthermore, given the high correlation between satisfaction and acceptance, the differences found between these two variables may be interpreted as chance findings. Lastly, this study's sufficient sample size and power to detect differences likely means that there was not a high amount of variance to account for in these groups. It is unlikely that a larger sample would produce significantly different effects.

One explanation for the partially significant results is the type of items used to measure the outcome variables. The satisfaction scale consisted of items such as, "I would be satisfied receiving this performance feedback form", and may have been relatively easy to endorse due to their affective nature. Conversely, the acceptance scale items were worded in a more knowledge-based manner, such as "This form could reflect feedback that is based on true class performance". Participants may have believed they were less experienced or knowledgeable to make distinctions between forms, thus making them more difficult to endorse and/or subject to measurement error (Raykov & Marcoulides, 2010).

Despite marginally significant results, this study demonstrates a need for clear performance expectations when designing and implementing feedback forms. The practical implications of these findings should encourage feedback form developers to detail the expectations of performance clearly and carefully, as these steps can account for a ratee's satisfaction with the form they receive. Given the relatively minor changes that were made to transform the feedback forms from the low clarity condition to one of high clarity, this extra level of detail could be easily incorporated into form development. From a research perspective, further investigation of clarity's impact on actual behavioral change would be valuable. While

the present study used attitudinal outcomes as a proxy of behavioral change likelihood, testing the relationship between a form's clarity of expectations and actual behavioral change would arguably be more compelling (Stone & Stone, 1985; Thurston et al, 2009). Further exploration of this idea is discussed as a future direction below.

Interaction. It was hypothesized that there would be an interaction between rating format and expectation clarity, such that participants would prefer numerically based feedback forms in low clarity conditions, but this preference would lesson in high clarity conditions. Results did not support this hypothesis. These findings indicate that the interaction between rating format and expectation clarity is not a meaningful source of information for understanding ratee attitudes towards feedback forms. As this was an exploratory hypothesis, this non-significant result may inform future studies interested in similar interactions.

Fairness as a Mediator. Fairness was hypothesized to mediate the relationship between both independent variables of rating format and expectation clarity, and both dependent measures of satisfaction and acceptance. The results did not indicate a possible mediation for any of the proposed relationships (Baron et al., 1986), including the marginally significant relationship between form clarity and ratee satisfaction. Mediation was unable to be fully tested, as neither of the independent variables (rating format and clarity of expectations) were significantly related to the proposed mediating variable of fairness. Without any connection between the independent variables and the mediator, there was no opportunity for the mediator to explain the relationship between the independent variables and any outcome.

There are several ways to interpret these results. Fairness is a complex and layered attitude, with a strong social component (Colquitt et al., 2001). It is possible that fairness was not salient to participants in this study because they were not asked to reflect on many of the

variables that shape this attitude, such as their relationship with the instructor that would be using the form or their past experiences with feedback from the university. Implications of the rater-ratee relationship will be discussed in the limitations section below.

Relatedly, it is possible that rating format and expectation clarity were simply not relevant enough to fairness. Fairness is most commonly correlated with other job attitudes and social interactions, and while the antecedents of justice remain less understood, variables such as voice, trust, and respect have been shown to shape fairness perceptions (Greenberg et al., 2013). This study's independent variables did not tap any of these key fairness-related aspects. Therefore, focusing on the relatively narrow and minor aspects of a feedback form may not have been enough to evoke the affective reactions needed to form a fairness attitude. This would explain why fairness did not impact the relationships in this study, while still allowing for the possibility that it is an important part of the feedback process in general. Therefore, although the results of this non-significant mediation indicate that fairness does not have a relationship with rating format or expectation clarity, the scope of this study was not broad enough to infer that justice has no place in the general field of performance feedback research.

Limitations

Like most, this study was not without its limitations. The field of performance feedback is extremely broad, and it was necessary to narrow the scope of this research question to a relatively small area of the field. Therefore, related variables of interest that may have been useful for answering the research question at hand were excluded from analysis. These issues, as well as other drawbacks, will be explored in detail below.

Policy-capturing. One of the main issues with using policy-capturing method is the compromise between sample size and power. As policy-capturing results are based on each individual's responses, the power of the individual analysis and the significant of regression weights depends on the amount of opportunities that a participant has to make a judgment (Karren & Barringer, 2002). This was acknowledged in my study design to an extent, as participants rated two sets of the same four forms. Ideally, though, participants would have rated five to ten sets of the same four forms (Cooksey, 1996). This would have required each participant to rate 20 to 40 forms, not including any distractor forms that would be included. However, the risk of boredom and reduced reliability are strongly implied with increased scenarios, particularly when these scenarios appear to be repetitive or similar (Graham & Cable, 2001). This trade-off was carefully considered, and given the form's limited variance in appearances and the sample of undergraduate students, the likelihood of boredom resulting in response invariance seemed high. Therefore, effort was taken to reduce the amount of forms with the knowledge that significance levels may be impacted. Future studies with the ability to offer monetary incentives, or studies employing a different sample, may be able to include more forms without significantly increasing the likelihood of response invariance or boredom.

Rater-ratee relationship. In actual practice, the social component of the rater-ratee relationship can play a significant role in performance feedback and appraisal attitudes (Pichler, 2012). Previous experiences with the supervisor who manages the feedback, the power distance between supervisors and subordinates, and the supervisor's feedback style can influence a ratee's likelihood of using the feedback (Pichler, 2012). These aspects were not addressed within the present study, as participants were only rating the feedback form. It was also outside of this

study's scope to include a performance feedback content component. However, given the strong social normative aspect of attitude formation (Ajzen, 1991), it is possible that supervisor relationship (or, in this case, instructor relationship) could impact attitudes towards feedback forms even without considering the feedback content. If this study were to measure reactions to feedback content as well as the feedback form, then measuring the ratee-rater relationship could prove valuable. Within the present study, the inclusion of scale items to measure the student-teacher relationship may have accounted for variance in attitudes towards feedback forms.

Narrative Feedback. There is an important distinction between qualitative and narrative feedback. Narrative feedback goes beyond text rating formats to include a supervisor's qualitative explanations of why an individual received certain feedback, how to better meet expectations, and any other relevant information. Raters using narrative feedback can theoretically provide as much (or as little) as they want, though they are not required to receive training on how to present their feedback (Kaiser & Kaplan, 2005). Historically, the research on feedback rating format has been focused on different types of *quantitative* feedback (Brutus, 2010; Ilgen, Barnes-Farrell, & McKellin, 1993; Landy & Farr, 1980; Sulsky & Keown, 1998), with little focus on how to improve or incorporate qualitative information. This is interesting when we consider that most performance feedback systems include, in some way, a qualitative component (Brutus, 2010; Rose & Walsh, 2004; Smither & Walker, 2004; Timmreck & Bracken, 1995; Woods, Sciarini, & Breiter, 1998). Brutus (2010) presented a theoretical review which argued for the importance of more research on qualitative feedback, and accurately predicted a rising interest in this research. This recent increase in studying narrative feedback can be attributed to several factors. Technological advances have allowed organizations to report and compile qualitative feedback in a more organized, efficient, and timely manner than when

compared to traditional paper-based methods. The option of anonymity via technology also opens up the possibility for qualitative feedback to be more honest, direct, and useful (Christianson-DeMay, Chandonnet, Rasinowich, and Fenlason, 2006). While the research on narrative feedback remains mixed regarding its benefits or advantages over traditional feedback formats, it may be possible that attitudes towards the feedback form differ as a result of this format type.

Future Directions

The results of this study, coupled with an understanding of its limitations, can help shape the future directions of performance feedback research. In this study, it was theorized that ratee reactions to feedback forms are important because they will shape behavioral changes that are made in response to the feedback content. While the connection between reactions to feedback and the likelihood of using that feedback have found support in previous research (Gosselin et al., 1997; Matsumara et al., 2004), reactions to feedback forms has received less attention. Testing this theory would provide more support for the importance of feedback form appeal. One ideal way to test this hypothesis would be in a field study, using the real performance feedback that is presented to employees but varying the feedback forms they receive it in. Follow-up measures of the employees' intention to use the feedback could identify differences by form condition, provided that the type of feedback (positive or negative) was controlled for. This would provide an opportunity to re-measure fairness as a mediating variable under more appropriate conditions, as well as measure the effect of supervisor relationship.

More broadly, in order to determine if a feedback form can influence ratee reactions, further research should be conducted on the most beneficial kinds of form manipulation. As the present study has demonstrated, rating formats of text versus numerical do not impact a ratee's

preference for the type of form they receive. However, the clarity of the performance expectations did account for a marginal amount of variance in ratee satisfaction. These variables represent just two of the many feedback form aspects that could be meaningfully contributing to reactions. Other feedback form features that merit investigation are different types of rating formats beyond the ones measured here (e.g. using colors or graphics instead of words or numbers), the type of language used (e.g. second or third person voice), the presence or absence of comparative information (i.e. how many other people received a similar feedback rating), and the inclusion of social cues (e.g. listing the names of supervisors, including the organization's mission statement, etc).

CONCLUSION

Performance feedback is one of the most highly researched areas in industrial/organizational psychology, but it remains far from being fully understood. Ratee attitudes towards their feedback forms can shape their attitudes about the feedback content in general. These attitudinal outcomes are crucial to understand, as they have been linked with changes to behavior and performance in other research (Azjen, 1991; Gosselin et al., 1997; Matsumara et al., 2004). This research demonstrated the importance of a feedback form's expectation clarity on ratee satisfaction with the feedback form, and connects these findings with larger implications of behavioral change. As one of the principle tenets of industrial/organizational psychology is to improve the performance of employees, this type of insight is not only valuable, but also crucial for furthering the field.

REFERENCES

Aiman-Smith, L., Scullen, S., & Barr, S. (2002). Conducting studies of decision making in organizational contexts: A tutorial for policy-capturing and other regression-based techniques. *Organizational Research Methods*, *5*(4), 388-414.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179-211.

Ajzen, I., & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology*, *22*(5), 453-474.

Alvero, A., Bucklin, B., & Austin, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985-1998). *Journal of Organizational Behavior Management*, *21*(1), 3-29.

Ammons, R. B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formulation. *The Journal of General Psychology*, *54*(2), 279-299.

Anseel, F., & Lievens, F. (2009). The mediating role of feedback acceptance in the relationship between feedback and attitudinal and performance outcomes. *International Journal of Selection and Assessment*, *17*(4), 362-376.

Arps, G. F. (1920). Work with knowledge of results versus work without knowledge of results: Awareness and partial awareness as factors conditioning efficiency. *Psychological Monographs*, *28*(3), i.

Ashford, S. J., & Cummings, L. L. (1983). Feedback as an individual resource: Personal strategies of creating information. *Organizational Behavior and Human Performance*, *32*(3), 370-398.

Avnet, T., & Higgins, E. (2003). Locomotion, assessment, and regulatory fit: Value transfer from "how" to "what". *Journal of Experimental Social Psychology*, *39*(5), 525-530.

Babcock, R. A., Sulzer-Azaroff, B., Sanderson, M., & Scibak, J. (1992). Increasing nurses' use of feedback to promote infection-control practices in a head-injury treatment center. *Journal of Applied Behavior Analysis, 25*(3), 621-627.

Bacal, R. (2003). *The manager's guide to performance reviews*. McGraw Hill Professional.

Balcazar, F. E., Shupert, M. K., Daniels, A. C., Mawhinney, T. C., & Hopkins, B. L. (1989). An objective review and analysis of ten years of publication in the Journal of Organizational Behavior Management. *Journal of Organizational Behavior Management*, *10*(1), 7-37.

Baron, R., and Kenny, D. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Behrman, D.N., Bigoness, W.J. and Perreault, W.D. (1982). Sources of job-related ambiguity and their consequences upon salespersons' job satisfaction and performance. *Management Science, 27*(1982), 1246-1260.

Bies, R. J., & Moag, J. S. (1986). Interactional justice: Communication criteria of fairness. *Research on Negotiation in Organizations*, *1*(1), 43-55.

Blau, G. (1999). Testing the longitudinal impact of work variables and performance appraisal satisfaction on subsequent overall job satisfaction. *Human Relations*, *52*(8), 1099-1113.

Bobko, P., & Coella, A. (1994). Employee reactions to performance standards: A review and research propositions. *Personnel Psychology*, *47*(1), 1-29.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. *Personnel selection in organizations,* San Francisco: Jossey-Bass.

Boswell, W. R., & Boudreau, J. W. (2002). Separating the developmental and evaluative performance appraisal uses. *Journal of Business and Psychology*, *16*(3), 391-412.

Brand, J. E. (1905). The effect of verbal suggestion upon the estimation of linear magnitudes. *Psychological Review, 12*(1), 41-49.

Bretz, R. D., Milkovich, G. T., & Read, W. (1992). The current state of performance appraisal research and practice: Concerns, directions, and implications. *Journal of Management*, *18*(2), 321-352.

Brown, M., Hyatt, D., & Benson, J. (2010). Consequences of the performance appraisal experience. *Personnel Review*, *39*(3), 375-396.

Brutus, S. (2010). Words versus numbers: A theoretical exploration of giving and receiving narrative comments in performance appraisal. *Human Resource Management Review*, *20*(2), 144-157.

Byrne, Z. S. (2005). Fairness reduces the negative effects of organizational politics on turnover intentions, citizenship behavior and job performance. *Journal of Business and Psychology*, *20*(2), 175-200.

Cable, D. M., & Judge, T. A. (1994). Pay preferences and job search decisions: A person-organization fit perspective. *Personnel Psychology*, *47*(2), 317-348.

Calpin, J. P., Edelstein, B., & Redmon, W. K. (1988). Performance feedback and goal setting to improve mental health center staff productivity. *Journal of Organizational Behavior Management*, *9*(2), 35-58.

Cardy, R. L., & Keefe, T. J. (1994). Observational purpose and evaluative articulation in frame-of-reference training: The effects of alternative processing modes on rating accuracy. *Organizational Behavior and Human Decision Processes*, *57*(3), 338-357.

Cawley, B. D., Keeping, L. M., & Levy, P. E. (1998). Participation in the performance appraisal process and employee reactions: A meta-analytic review of field investigations. *Journal of Applied Psychology*, *83*(4), 615-633.

Cesario, J., Higgins, E. T., & Scholer, A. A. (2008). Regulatory fit and persuasion: Basic principles and remaining questions. *Social and Personality Psychology Compass*, *2*(1), 444-463.

Chambers, T. M., Anderson, P. F., & Dunlap, B. J. (1986). Preferences, intentions, and behavior of organizational buyers under different reward conditions. *Journal of Business Research*, *14*(6), 533-547.

Christianson-DeMay, C., Chandonnet, A., Rasinowich, C., & Fenlason, K. J. (2006). Application of an automated content analysis process to multisource comments. *Paper presented at the 21st Annual Meeting of the Society for Industrial and Organizational Psychology, Dallas, TX.*

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Psychology Press.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, *1*(3), 98-101.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression-correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *86*(2), 278-321.

Colquitt, J.A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology, 86*(3), 386-400.

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, *86*(3), 425.

Cooksey, R. (1996). *Judgment analysis: Theory, methods, and applications.* San Diego, CA: Academic Press.

Crawley, S. L. (1926). An experimental investigation of recovery from work. *Archives of Psychology.*

Culbertson, S., Henning, J., & Payne, S. (2013). Performance appraisal satisfaction: The role of feedback and goal orientation. *Journal of Personnel Psychology, 12*(4), 189-195.

Dai, G., De Meuse, K. P., & Peterson, C. (2010). Impact of multi-source feedback on leadership competency development: A longitudinal field study. *Journal Of Managerial Issues*, *22*(2), 197-219.

DeNisi, A. & Pritchard, R. (2006). Performance appraisal, performance management, and improving individual performance: A motivational framework. *Management and Organization Review, 2*(2), 253-277.

Deutsch, M. (1985). *Distributive justice: A social-psychological perspective*. New Haven: Yale University Press.

DeVries, J., Burnette, M., & Redmon, W. (1991). AIDS prevention: Improving nurses' compliance with glove wearing through performance feedback. *Journal of Applied Behavior Analysis*, *24*(4), 705-711.

Dusterhoff, C., Cunningham, J., & MacGregor, J. (2013). The effects of performance rating, Leader–Member Exchange, perceived utility, and organizational justice on performance appraisal satisfaction: Applying a moral judgment perspective. *Journal of Business Ethics, 119*(2), 265-273.

Eagly, A. H,, & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt.

Elliot, E., & Dweck, C. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology, 54*(1), 5-12.

Evans, C., & Waring, M. (2011). Student teacher assessment feedback preferences: The influence of cognitive styles and gender. *Learning And Individual Differences*, *21*(3), 271-280.

Feys, M., Anseel, F., & Wille, B. (2011). Improving feedback reports: the role of procedural information and information specificity. *Academy of Management Learning & Education*, *10*(4), 661-681.

Fletcher, C. (1997). *Appraisal: Routes to improved performance*. London: Chartered Institute of Personnel & Development.

Fletcher, C. (2001). Performance appraisal and management: The developing research agenda. *Journal of Occupational and Organizational Psychology*, *74*(4), 473-487.

Fox, C. J., & Sulzer-Azaroff, B. (1989). The effectiveness of two different sources of feedback on staff teaching of fire evacuation skills. *Journal of Organizational Behavior Management, 10*(2), 19-35.

Gaetani, J., Hoxeng, D., & Austin, J. (1985). Engineering compensation systems: Effects of commissioned versus wage payment. *Journal of Organizational Behavior Management, 7*(1-2)*,* 51-63.

Gates, A. I. (1922). *Recitation as a factor in memorizing* (No. 40). Science Press.

Giles, W. F., & Mossholder, K. W. (1990). Employee reactions to contextual and session components of performance appraisal. *Journal of Applied Psychology, 75*(4)*,* 371-377.

Gosselin, A., Werner, J., & Halle, N. (1997). Ratee preferences concerning performance management and appraisal. *Human Resource Development Quarterly, 8*(4), 315-333.

Graham, M. E., & Cable, D. M. (2001). Consideration of the incomplete block design for policy-capturing research. *Organizational Research Methods, 4*(1), 26-45.

Green, S., Salkind, N. J., & Akey, T. M. (2003). *For Windows: Analyzing and Understanding Data.* Prentice Hall.

Greenberg, J. (1986). Determinants of perceived fairness of performance evaluations. *Journal of Applied Psychology, 71*(2), 340-342.

Greenberg, J., & Colquitt, J. A. (Eds.). (2013). *Handbook of organizational justice*. Psychology Press.

Greenberg, J., & Greenberg, J. (1993). The social side of fairness: Interpersonal and informational classes of organizational justice. *Justice in the workplace: Approaching fairness in human resource management,* Lawrence Erlbaum Associates, Hillsdale, NJ.

Greguras, G. J., Ford, J. M., & Brutus, S. (2003). Manager attention to multisource feedback. *Journal Of Management Development*, *22*(4), 345-361.

Hancock, G. R., & Mueller, R. O. (Eds.). (2010). *The reviewer's guide to quantitative methods in the social sciences*. Routledge.

Hair, J., Anderson, R., Tatham, R., & Black, W. (1992). *Multivariate data analysis with readings*. New York: Macmillan.

Hawkins, A., Burgio, L., Lanford, A., & Engel, B. (1992). The effects of verbal and written supervisory feedback on staff compliance with assigned prompted voiding in a nursing home. *Journal of Organizational Behavior Management*, *13*(1), 137-150.

Hedge, J. W., & Borman, W. C. (1995). Changing conceptions and practices in performance appraisal. *The changing nature of work,* San Francisco: Jossey-Bass.

Heneman III, H., & Milanowski, A. (2003). Continuing assessment of teacher actions to a standards-based teacher evaluation system. *Journal of Personnel Evaluation in Education*, *17*(2), 173-195.

Hinton, P., Brownlow, C., & McMurray, I. (2004). *SPSS Explained*. Routledge.

Hitt, M., & Barr, S. (1989). Managerial selection decision models: Examination of configural cue processing. *Journal of Applied Psychology*, *74*(1), 53-61.

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology*, *27*(1), 99-114.

Ilgen, D. R., Barnes-Farrell, J. L., & McKellin, D. B. (1993). Performance appraisal process research in the 1980s: What has it contributed to appraisals in use? *Organizational Behavior and Human Decision Processes, 54*(3), 321−368.

Ilgen, D. R., & Davis, C. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology*, *49*(3), 550-565.

Ilgen, D. R., & Favero, J. (1985). Limits in generalization from psychological research to performance appraisal process. *Academy of Management Review, 10*(2), 311-321.

Ilgen, D. R., & Feldman, J. (1983). Performance appraisal: A process focus. *Research in organizational behavior*, Greenwich, CT: JAI Press.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology, 64*(4), 349–371.

Ilgen, D. R., & Hollenback, J. H. (1977). The role of job satisfaction in absence behavior. *Organizational Behavior and Human Performance, 19*(1)*,* 148-161.

Jawahar, I. M. (2005). Do raters consider the influence of system factors on observed performance when evaluating performance? Evidence from three experiments. *Group & Organization Management, 30*(1)*,* 6-41.

Jawahar, I.M. (2006). An investigation of potential consequences of satisfaction with appraisal feedback. *Journal of Leadership and Organizational Studies, 13*(2), 14-28.

Jawahar, I. M. (2007). The influence of perceptions of fairness on performance appraisal reactions. *Journal of Labor Research, 28*(4)*,* 735-754.

Judd, C. H. (1905). Practice without knowledge of results. *The Psychological Review: Monograph Supplements*.

Kaiser, R. B., & Kaplan, R. E. (2005). Overlooking overkill? Beyond the 1-to-5 rating. *Human Resource Planning, 28*(1), 7-11.

Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations*.

Karren, R., & Barringer, M. (2002). A review and analysis of the policy-capturing methodology in organizational research: Guidelines for research and practice. *Organizational Research Methods, 5*(4), 337-361.

Keeping, L., & Levy, P. (2000). Performance appraisal reactions: Measurement, modeling, and method bias. *Journal of Applied Psychology, 85*(5), 708–723.

Kline, R. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. American Psychological Association.

Kline, T., & Sulsky, L. (2009). Measurement and assessment issues in performance appraisal. *Canadian Psychological Association, 50*(3), 161-171.

Kluger, A. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119(2),* 254-284.

Krivoshey, R. (1975). *Instructions, verdicts, and judicial behavior.* Taylor & Francis.

Kulas, J. T., & Finkelstein, L. M. (2007). Content and reliability of discrepancy-defined self-awareness in multisource feedback. *Organizational Research Methods*, *10*(3), 502-522.

Kuvaas, B. (2006). Performance appraisal satisfaction and employee outcomes: mediating and moderating roles of work motivation. *The International Journal of Human Resource Management, 17*(3), 504-522.

Lam, S. S., Yik, M. S., & Schaubroeck, J. (2002). Responses to formal performance appraisal feedback: The role of negative affectivity. *Journal of Applied Psychology, 87*(1)*,* 192–201.

Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin, 87*(1), 72–107.

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. Routledge.

Latham, G. P.. & Wexley, K. N. (1981). *Increasing productivity through performance appraisal*. Reading, MA: Addison-Wesley.

Latham, G. P., & Yukl, G. A. (1975). A review of research on the application of goal setting in organizations. *Academy of Management Journal*, *18*(1), 824-845.

Latham, G. P., & Yukl, G. A. (1976). Effects of assigned and participative goal setting on performance and job satisfaction. *Journal of Applied Psychology*, *61*(2), 166.

Locke, E. (1968). Towards a theory of task motivation incentives. *Organizational Behavior and Human Performance*, 3(1), 157-189.

Locke, E., & Latham, G. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Mamatoglu, N. (2008). Effects on organizational context (culture and climate) from implementing a 360-degree feedback system: The case of Arcelik. *European Journal Of Work And Organizational Psychology*, *17*(4), 426-449.

Manshor, A., & Kamalanabhan, T. (2000). An examination of raters' and ratees' preferences in process and feedback in performance appraisal. *Psychological Reports*, *86*(1), 203-214.

Matsumura, S., & Hann, G. (2004). Computer anxiety and students' preferred feedback methods in EFL writing. *The Modern Language Journal*, *88*(3), 403-415.

Muchinsky, P. M. (2012). *Psychology applied to work.* Summerfield, NC: Hypergraphic Press.

Murphy, K. R., & Cleveland, J. N. (1991). *Performance appraisal: An organizational perspective*. Boston: Allyn & Bacon.

Murphy, K. R., & Cleveland, J. N. (1995). *Understanding performance appraisal: Social, organizational, and goal-based perspectives*. Thousand Oaks: Sage.

Murphy, K. R., & Myors, B. (2004). *Statistical power analysis*. Mahwah, New Jersey.

Neter, J., Kutner, M., Nachtsheim, C., & Wasserman,W. (1996*). Applied linear regression models.* Chicago: Irwin.

Nordstrom, R., Hall, R., Lorenzi, P., & Delquadri, J. (1988). Organizational behavior modification in the public sector: Three field experiments. *Journal of Organizational Behavior Management, 9*(2)*,* 91-112.

Pettijohn, C. E., Pettijohn, L. S., & d'Amico, M. (2001). Characteristics of performance appraisals and their impact on sales force satisfaction. *Human Resource Development Quarterly, 12*(2)*,* 127–146.

Petty, M., Singleton, B., & Connell, D. (1992). An experimental evaluation of an organization incentive plan in the electric utility industry. *Journal of Applied Psychology, 77*(4), 427-436.

Pichler, S. (2012). The social context of performance appraisal and appraisal reactions: A meta-analysis. *Human Resource Management, 51*(5), 709–732.

Porterfield, J., Evans, G., & Blunden, R. (1985). Involving families and staff in service improvement. *Journal of Organizational Behavior Management, 7*(1-2), 117-133.

Posthuma, R., & Campion, M. (2008). Twenty best practices for job employee performance reviews. *Compensation and Benefits Review, 40(1),* 47–55.

Raykov, T, & Marcoulides, G. (2010). *Introduction to psychometric theory*. Taylor & Francis.

Roberson, Q. M., & Stewart, M. M. (2006). Understanding the motivational effects of procedural and informational justice in feedback processes. *British Journal of Psychology*, *97*(3), 281-298.

Rose, D.S., & Walsh, A.B. (2004). Current trends in 360° feedback. *Technical report #8251*. Data Driven Decisions, Inc.

Russell, J. S., & Goode, D. L. (1988). An analysis of managers' reactions to their own performance appraisal feedback. *Journal of Applied Psychology, 73*(1), 63-67.

Sanchez, J., & Levine, E. (1989). Determining important tasks within jobs: A policy capturing approach. *Journal of Applied Psychology*, *74*(2), 336-342.

Seifert, C. F., & Yukl, G. (2010). Effects of repeated multi-source feedback on the influence behavior and effectiveness of managers: A field experiment. *The Leadership Quarterly*, *21*(5), 856-866.

Sherer, P., Schwab, D., & Heneman, H. G. (1987). Managerial salary-raise decisions: A policy-capturing approach. *Personnel Psychology*, *40*(1), 27-38.

Shimp, T. A., Hyatt, E. M., & Snyder, D. J. (1991). A critical appraisal of demand artifacts in consumer research. *Journal of Consumer Research, 18*(1), 273-283.

Siero, S., Boon M., Kok, G., & Siero, F. (1989). Modification of driving behavior in a large transport organization: A field experiment. *Journal of Applied Psychology*, *74*(3), 417-423.

Smither, J. W. (1998). *Lessons learned: Research implications for performance appraisal and management*. San Francisco: Jossey- Bass.

Smither, J. W., & Walker, A. G. (2004). Are the characteristics of narrative comments related to improvement in multirater feedback over time? *Journal of Applied Psychology, 89*(5), 575−581.

Spencer, L. T. (1923). Minor studies from the psychological laboratory of Yale University. *American Journal of Psychology, 34*(1), 107-111.

Stone-Romero, F. (1994). Construct validity issues in organizational behavior research. *Organizational behavior: The state of the science*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Stone, D. L., & Stone, E. F. (1985). The effects of feedback consistency and feedback favorability on self-perceived task competence and perceived feedback accuracy. *Organizational Behavior and Human Decision Processes*,*36*(2), 167-185.

Sulsky, L. M., & Keown, J. L. (1998). Performance appraisal in the changing world of work: Implications for the meaning and measurement of work performance. *Canadian Psychology, 39*(1-2)*,* 52−59.

Thibaut, J. & Walker, L. (1975). *Procedural justice: A psychological analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Timmreck, C.W., & Bracken, D.W. (1995). *Upward feedback in the trenches: Challenges and realities.* In *10th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.*

Thoms, P., McMasters, R., Roberts, M., & Dombkowski, D. (1999). Resume characteristics as predictors of an invitation to interview. *Journal of Business and Psychology*, *13*(3), 339-356.

Thorndike, E. L. (1927). The law of effect. *American Journal of Psychology*, *39*(1), 212-222.

Thorndike, R. L. (1949). *Personnel selection*. New York: Wiley.

Thurston, P., & McNall, L. (2009). Justice perceptions of performance appraisal practices. *Journal of Managerial Psychology, 25*(3), 201-228.

Toegel, G., & Conger, J. A. (2003). 360-degree assessment: Time for reinvention. *Academy Of Management Learning & Education*, *2*(3), 297-311.

Van Vianen, A.E.M., Taris, R., Scholten, E. and Schinkel, S. (2004) Perceived fairness in personnel selection: Determinants and outcomes in different stages of the assessment procedure. *International Journal of Selection and Assessment, 12*(1-2)*,* 149–159.

Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.

Waters, R. H. (1933). The specificity of knowledge of results and improvement. *Psychological Bulletin, 30*(1), 673.

Woods, R. H., Sciarini, M., & Breiter, D. (1998). Performance appraisal in hotels. *The Cornell Hotel and Restaurant Administration Quarterly, 39*(2), 25−29.

Wright, W. R. (1906). Some effects of incentive on work and fatigue. *Psychological Review, 13*(1), 23-34.

Zedeck, S., & Kafry, D. (1977). Capturing rater policies for processing evaluation data. *Organizational Behavior and Human Performance*, *18*(2), 269-294.

APPENDIX

Appendix A:
Instruction Set

Instructions

The Psychology Department at Colorado State University is interested in creating a form for instructors to provide students with feedback on individual class performance. All programs and colleges within the university will use the same feedback form. We are asking for student input on these feedback forms. Your opinion will directly contribute to the final product.

You will be shown sample feedback forms and asked to provide your reaction to these forms. You will also be asked to rank order your preference of the forms. Please be prepared to explain which aspects of each form are impacting your decision for preference.

Appendix B:
Form for Text/High Clarity Condition (Set 1)

PERFORMANCE FEEDBACK

I. **STUDENT INFORMATION**

STUDENT NAME	COURSE CODE

II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER

III. PRIMARY PURPOSE

- Provide information on student's class performance.

IV. FEEDBACK

1. EXAMS (Forty percent of total grade)

RATING: □ POOR □BELOW AVERAGE □AVERAGE □ABOVE AVERAGE □EXCEPTIONAL

- Exam performance is based on your average exam score.
- Possible performance levels range from receiving no points on any exams to receiving full points on all exams to date.
- A rating of "poor" indicates an average exam score of 0%-45%.
- A rating of "average" indicates an average exam score of 70%-79%.
- A rating of "exceptional" indicates an average exam score of 90%-100%.

Instructor comments:

2. QUIZZES (Twenty percent of total grade)

RATING: □ POOR □BELOW AVERAGE □AVERAGE □ABOVE AVERAGE □EXCEPTIONAL

- Quiz performance is based on participation.
- Possible performance levels range from zero quiz participation to participation in all quizzes.
- A rating of "poor" indicates no more than 40% of quizzes have been submitted.
- A rating of "average" indicates no more than 75% of quizzes have been submitted.
- A rating of "exceptional" indicates at least 90% of quizzes have been submitted.

Instructor comments:

3. PROJECTS (Thirty five percent of total grade)

RATING: □ POOR □BELOW AVERAGE □AVERAGE □ABOVE AVERAGE □EXCEPTIONAL

- Project performance is based on your average project grade.
- Possible performance levels range from receiving no points on any projects to receiving full points on all projects to date.
- A rating of "poor" indicates an average project score of 0%-45%.
- A rating of "average" indicates an average project score of 70%-79%.
- A rating of "exceptional" indicates an average project score of 90%-100%.

Instructor comments:

Appendix C:
Form for Text/High Clarity Condition (Set 2)

PERFORMANCE FEEDBACK

I. **STUDENT INFORMATION**

STUDENT NAME	COURSE CODE

II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER

III. PRIMARY PURPOSE

- Provide information on student's class performance.

IV. FEEDBACK

1. EXAMS (Forty percent of total grade)

RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL

* Exam performance is based on your average exam score.
* Possible performance levels range from receiving no points on any exams to receiving full points on all exams to date.
* A rating of "poor" indicates an average exam score of 0%-45%.
* A rating of "average" indicates an average exam score of 70%-79%.
* A rating of "exceptional" indicates an average exam score of 90%-100%.

Instructor comments:

2. QUIZZES (Twenty percent of total grade)

RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL

* Quiz performance is based on participation.
* Possible performance levels range from zero quiz participation to participation in all quizzes.
* A rating of "poor" indicates no more than 40% of quizzes have been submitted.
* A rating of "average" indicates no more than 75% of quizzes have been submitted.
* A rating of "exceptional" indicates at least 90% of quizzes have been submitted.

Instructor comments:

3. PROJECTS (Thirty five percent of total grade)

RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL

- Project performance is based on your average project grade.
- Possible performance levels range from receiving no points on any projects to receiving full points on all projects to date.
- A rating of "poor" indicates an average project score of 0%-45%.
- A rating of "average" indicates an average project score of 70%-79%.
- A rating of "exceptional" indicates an average project score of 90%-100%.

Instructor comments:

Appendix D:
Form for Text/Low Clarity Condition (Set 1)

PERFORMANCE FEEDBACK

I. **STUDENT INFORMATION**

STUDENT NAME	COURSE CODE

II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER

III. PRIMARY PURPOSE

- Provide information on student's class performance.

IV. FEEDBACK

1. EXAMS (Forty percent of total grade)

RATING: □ POOR □BELOW AVERAGE □AVERAGE □ABOVE AVERAGE □EXCEPTIONAL

- Exams are an important part of your course grade.
- Possible performance levels range from "poor" to "exceptional".
- A rating of "poor" indicates very low exam scores.
- A rating of "average" indicates passing exam scores.
- A rating of "exceptional" indicates high exam scores.

Instructor comments:

2. QUIZZES (Twenty percent of total grade)

RATING: □ POOR □BELOW AVERAGE □AVERAGE □ABOVE AVERAGE □EXCEPTIONAL

- Quizzes are an important part of your course grade.
- A rating of "poor" indicates that few or no quizzes have been submitted.
- A rating of "average" indicates that most quizzes have been submitted.
- A rating of "exceptional" indicates that mostly all quizzes have been submitted.

Instructor comments:

3. PROJECTS (Thirty five percent of total grade)

RATING: □ POOR □BELOW AVERAGE □AVERAGE □ABOVE AVERAGE □EXCEPTIONAL

* Projects are an important part of your course grade.
* A rating of "poor" indicates very low project scores.
* A rating of "average" indicates passing project scores.
* A rating of "exceptional" indicates high project scores.

Instructor comments:

Appendix E:
Form for Text/Low Clarity Condition (Set 2)

PERFORMANCE FEEDBACK

I. **STUDENT INFORMATION**

STUDENT NAME	COURSE CODE

II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER

III. PRIMARY PURPOSE

- Provide information on student's class performance.

IV. FEEDBACK

1. EXAMS (Forty percent of total grade)

RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL

- Exams are an important part of your course grade.
- Possible performance levels range from "poor" to "exceptional".
- A rating of "poor" indicates very low exam scores.
- A rating of "average" indicates passing exam scores.
- A rating of "exceptional" indicates high exam scores.

Instructor comments:

2. QUIZZES (Twenty percent of total grade)

RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL

- Quizzes are an important part of your course grade.
- A rating of "poor" indicates that few or no quizzes have been submitted.
- A rating of "average" indicates that most quizzes have been submitted.
- A rating of "exceptional" indicates that mostly all quizzes have been submitted.

Instructor comments:

3. PROJECTS (Thirty five percent of total grade)

RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL

* Projects are an important part of your course grade.
* A rating of "poor" indicates very low project scores.
* A rating of "average" indicates passing project scores.
* A rating of "exceptional" indicates high project scores.

Instructor comments:

Appendix F:
Form for Numerical/High Clarity Condition (Set 1)

PERFORMANCE FEEDBACK

I. **STUDENT INFORMATION**

STUDENT NAME	COURSE CODE

II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER

III. PRIMARY PURPOSE

- Provide information on student's class performance.

IV. FEEDBACK

1. EXAMS (40% of total grade)

RATING: □ **1** □**2** □**3** □**4** □**5**

- Exam performance is based on your average exam score.
- Possible performance levels range from receiving no points on any exams to receiving full points on all exams to date.
- A rating of "poor" indicates an average exam score of 0%-45%.
- A rating of "average" indicates an average exam score of 70%-79%.
- A rating of "exceptional" indicates an average exam score of 90%-100%.

Instructor comments:

2. QUIZZES (20% of total grade)

RATING: □ **1** □**2** □**3** □**4** □**5**

- Quiz performance is based on participation.
- Possible performance levels range from zero quiz participation to participation in all quizzes.
- A rating of "poor" indicates no more than 40% of quizzes have been submitted.
- A rating of "average" indicates no more than 75% of quizzes have been submitted.
- A rating of "exceptional" indicates at least 90% of quizzes have been submitted.

Instructor comments:

3. PROJECTS (35% of total grade)

RATING:	□ **1**	□**2**	□**3**	□**4**	□**5**

- Project performance is based on your average project grade.
- Possible performance levels range from receiving no points on any projects to receiving full points on all projects to date.
- A rating of "poor" indicates an average project score of 0%-45%.
- A rating of "average" indicates an average project score of 70%-79%.
- A rating of "exceptional" indicates an average project score of 90%-100%.

Instructor comments:

Appendix G:
Form for Numerical/High Clarity Condition (Set 2)

PERFORMANCE FEEDBACK

I. **STUDENT INFORMATION**

STUDENT NAME	COURSE CODE

II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER

III. PRIMARY PURPOSE

- Provide information on student's class performance.

IV. FEEDBACK

1. EXAMS (40% of total grade)

RATING: ¢ 1 ¢2 ¢3 ¢4 ¢5

* Exam performance is based on your average exam score.
* Possible performance levels range from receiving no points on any exams to receiving full points on all exams to date.
* A rating of "poor" indicates an average exam score of 0%-45%.
* A rating of "average" indicates an average exam score of 70%-79%.
* A rating of "exceptional" indicates an average exam score of 90%-100%.

Instructor comments:

2. QUIZZES (20% of total grade)

RATING: ¢ 1 ¢2 ¢3 ¢4 ¢5

* Quiz performance is based on participation.
* Possible performance levels range from zero quiz participation to participation in all quizzes.
* A rating of "poor" indicates no more than 40% of quizzes have been submitted.
* A rating of "average" indicates no more than 75% of quizzes have been submitted.
* A rating of "exceptional" indicates at least 90% of quizzes have been submitted.

Instructor comments:

3. PROJECTS (35% of total grade)

RATING:	¢ 1	¢2	¢3	¢4	¢5

* Project performance is based on your average project grade.
* Possible performance levels range from receiving no points on any projects to receiving full points on all projects to date.
* A rating of "poor" indicates an average project score of 0%-45%.
* A rating of "average" indicates an average project score of 70%-79%.
* A rating of "exceptional" indicates an average project score of 90%-100%.

Instructor comments:

Appendix H:
Form for Numerical/Low Clarity Condition (Set 1)

PERFORMANCE FEEDBACK

I. **STUDENT INFORMATION**

STUDENT NAME	COURSE CODE

II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER

III. PRIMARY PURPOSE

- Provide information on student's class performance.

IV. FEEDBACK

1. EXAMS (40% of total grade)

RATING: □ **1** □**2** □**3** □**4** □**5**

- Exams are an important part of your course grade.
- Possible performance levels range from "poor" to "exceptional".
- A rating of "poor" indicates very low exam scores.
- A rating of "average" indicates passing exam scores.
- A rating of "exceptional" indicates high exam scores.

Instructor comments:

2. QUIZZES (20% of total grade)

RATING: □ **1** □**2** □**3** □**4** □**5**

- Quizzes are an important part of your course grade.
- A rating of "poor" indicates that few or no quizzes have been submitted.
- A rating of "average" indicates that most quizzes have been submitted.
- A rating of "exceptional" indicates that mostly all quizzes have been submitted.

Instructor comments:

3. PROJECTS (35% of total grade)

RATING: □ **1** □**2** □**3** □**4** □**5**

* Projects are an important part of your course grade.
* A rating of "poor" indicates very low project scores.
* A rating of "average" indicates passing project scores.
* A rating of "exceptional" indicates high project scores.

Instructor comments:

Appendix I:
Form for Numerical/Low Clarity Condition (Set 2)

PERFORMANCE FEEDBACK

I. **STUDENT INFORMATION**

STUDENT NAME	COURSE CODE

II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER

III. PRIMARY PURPOSE

- Provide information on student's class performance.

IV. FEEDBACK

1. EXAMS (40% of total grade)

RATING: ¢ 1 ¢2 ¢3 ¢4 ¢5

- Exams are an important part of your course grade.
- Possible performance levels range from "poor" to "exceptional".
- A rating of "poor" indicates very low exam scores.
- A rating of "average" indicates passing exam scores.
- A rating of "exceptional" indicates high exam scores.

Instructor comments:

2. QUIZZES (20% of total grade)

RATING: ¢ 1 ¢2 ¢3 ¢4 ¢5

- Quizzes are an important part of your course grade.
- A rating of "poor" indicates that few or no quizzes have been submitted.
- A rating of "average" indicates that most quizzes have been submitted.
- A rating of "exceptional" indicates that mostly all quizzes have been submitted.

Instructor comments:

3. PROJECTS (35% of total grade)

RATING: ¢ 1 ¢2 ¢3 ¢4 ¢5

* Projects are an important part of your course grade.
* A rating of "poor" indicates very low project scores.
* A rating of "average" indicates passing project scores.
* A rating of "exceptional" indicates high project scores.

Instructor comments:

Appendix J:
Performance Feedback Form: Original Form (United States Air Force)

AFPAM36-2241 1 OCTOBER 2011

Figure 11.1. Sample AF Form 931, *Performance Feedback Worksheet (AB thru TSgt).*

PERFORMANCE FEEDBACK WORKSHEET *(AB thru TSgt)*

I. PERSONAL INFORMATION

NAME	GRADE	UNIT
TORES, RICHARDO J.	SSgt	AFSES/CM

II. TYPES OF FEEDBACK: [X] INITIAL [ ] MID-TERM [ ] RATEE REQUESTED [ ] RATER DIRECTED

III. PRIMARY DUTIES

- Outline specific duties completed to meet mission requirements
- The entries should include the most important duties and correspond to the job reflected in the EPR

IV. PERFORMANCE FEEDBACK

1. PRIMARY/ADDITIONAL DUTIES. Consider Adapting, Learning, Quality, Timeliness, Professional Growth, Communication Skills. (For SSgt/TSgt also consider Supervisory, Leadership and Technical Ability.)

[X] N/A Initial Feedback [ ] Does Not Meet [ ] Meets [ ] Above Average [ ] Clearly Exceeds

The following are some examples. Does the ratee: (1) display an ability to learn rapidly and adapt quickly to changing situations; (2) demonstrate accuracy, thoroughness, and orderliness in performing work assignments; (3) use systematic methods to accomplish more in less time; (4) actively support professional organizations; and (5) effectively communicate management decisions to achieve understanding and acceptance?

2. STANDARDS, CONDUCT, CHARACTER & MILITARY BEARING. Consider Dress & Appearance, Personal/Professional Conduct On/Off Duty. (For SSgt/TSgt also consider Enforcement of Standards and Customs & Courtesies.)

[X] N/A Initial Feedback [ ] Does Not Meet [ ] Meets [ ] Above Average [ ] Clearly Exceeds

The following are some examples. Does the ratee: (1) meet or exceed AF standards for dress and appearance; (2) project a positive military image according to AFI 36-2903; (3) achieve success when confronted with limited resources; (4) effectively overcome personal and organizational blocks to achieve results; and (5) support organizational, base and (or) community activities (duty or off-duty hours)?

3. FITNESS. Maintains Air Force Physical Fitness Standards.

[ ] Does Not Meet [X] Meets [ ] Exempt

Does the ratee participate in the AF physical training program and meet standards?

4. TRAINING REQUIREMENTS. Consider Upgrade, Ancillary, OJT, & Readiness. (For SSgt/TSgt also consider PME, Off-duty Education, Technical Growth, and Upgrade Training.)

[X] N/A Initial Feedback [ ] Does Not Meet [ ] Meets [ ] Above Average [ ] Clearly Exceeds

The following are some examples. Does the ratee: (1) complete training in the minimum time allowed; (2) accomplish readiness requirements in a timely manner; (3) have required PME for current grade completed; (4) translate innovative or better ways to do "things" into a plan of action; and (5) display a high level of technical competence?

5. TEAMWORK/FOLLOWERSHIP. Consider Team Building, Support of Team & Followership. (For SSgt/TSgt also consider Leadership, Team Accomplishments, Recognition/Reward Others.)

[X] N/A Initial Feedback [ ] Does Not Meet [ ] Meets [ ] Above Average [ ] Clearly Exceeds

The following are some examples. Is the ratee easy to work with? Does the ratee have a positive attitude? Does the ratee complain or foster team work to accomplish the task? Do the other unit, flight, or section personnel enjoy working with the ratee? Does the ratee display leadership traits appropriate to the situation? Does the ratee display the ability to simulate others?

6. OTHER COMMENTS. Consider Promotion, Future Duty/Assignment/Education Recommendations and Safety, Security & Human Relations.

[X] N/A Initial Feedback [ ] Does Not Meet [ ] Meets [ ] Above Average [ ] Clearly Exceeds

Consider future assignments that help the ratee achieve breadth of experience within their career field. Identify educational requirements needed to complete CCAF or higher degree. Ensure ratee exhibits safety and security practices daily to accomplish the mission. Also determine if the ratee works harmoniously and effectively with others.

AF FORM 931, 20080618 PREVIOUS EDITIONS ARE OBSOLETE PRIVACY ACT INFORMATION: The information in this form is FOR OFFICIAL USE ONLY. Protect IAW the Privacy Act of 1974.

Appendix K:
Distractor Form 1

PERFORMANCE FEEDBACK	
I. **STUDENT INFORMATION**	
STUDENT NAME	COURSE CODE
II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER	
III. PRIMARY PURPOSE - Provide information on student's class performance.	
IV. FEEDBACK	
1. EXAMS (Forty five percent of total grade)	
RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL	
Exams are an important part of your course grade. Possible performance levels range from "poor" to "exceptional". A rating of "poor" indicates very low exam scores. A rating of "average" indicates passing exam scores. A rating of "exceptional" indicates high exam scores.	
2. QUIZZES (Twenty percent of total grade)	
RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL	
Quizzes are an important part of your course grade. A rating of "poor" indicates that few or no quizzes have been submitted. A rating of "average" indicates that most quizzes have been submitted. A rating of "exceptional" indicates that mostly all quizzes have been submitted.	
3. PROJECTS (Thirty five percent of total grade)	
RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL	
Projects are an important part of your course grade. A rating of "poor" indicates very low project scores. A rating of "average" indicates passing project scores. A rating of "exceptional" indicates high project scores.	

Appendix L:
Distractor Form 2

PERFORMANCE FEEDBACK

I. **STUDENT INFORMATION**

STUDENT NAME	COURSE CODE

II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER

III. PRIMARY PURPOSE

- Provide information on student's class performance.

IV. FEEDBACK

1. EXAMS (45% of total grade)

RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL

Exam performance is based on your average exam score. Possible performance levels range from receiving no points on any exams to receiving full points on all exams to date. A rating of "poor" indicates an average exam score of 0%-45%. A rating of "average" indicates an average exam score of 70%-79%. A rating of "exceptional" indicates an average exam score of 90%-100%.

2. QUIZZES (20% of total grade)

RATING: ¢ 1 ¢2 ¢3 ¢4 ¢5

Quiz performance is based on participation. Possible performance levels range from zero quiz participation to participation in all quizzes. A rating of "poor" indicates no more than 40% of quizzes have been submitted. A rating of "average" indicates no more than 75% of quizzes have been submitted. A rating of "exceptional" indicates at least 90% of quizzes have been submitted.

3. PROJECTS (35% of total grade)

RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL

Project performance is based on your average project grade. Possible performance levels range from receiving no points on any projects to receiving full points on all projects to date. A rating of "poor" indicates an average project score of 0%-45%. A rating of "average" indicates an average project score of 70%-79%. A rating of "exceptional" indicates an average project score of 90%-100%.

Appendix M:
Distractor Form 3

PERFORMANCE FEEDBACK	
I. **STUDENT INFORMATION**	
STUDENT NAME	COURSE CODE
II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER	
III. PRIMARY PURPOSE - Provide information on student's class performance.	
IV. FEEDBACK	
1. EXAMS (Forty five percent of total grade)	
RATING: □ **POOR** □**AVERAGE** □**EXCEPTIONAL**	
Exams are an important part of your course grade. Possible performance levels range from "poor" to "exceptional". A rating of "poor" indicates very low exam scores. A rating of "average" indicates passing exam scores. A rating of "exceptional" indicates high exam scores.	
2. QUIZZES (Twenty percent of total grade)	
RATING: □ **POOR** □**AVERAGE** □**EXCEPTIONAL**	
Quiz performance is based on participation. Possible performance levels range from zero quiz participation to participation in all quizzes. A rating of "poor" indicates no more than 40% of quizzes have been submitted. A rating of "average" indicates no more than 75% of quizzes have been submitted. A rating of "exceptional" indicates at least 90% of quizzes have been submitted.	
3. PROJECTS (Thirty five percent of total grade)	
RATING: □ **POOR** □**AVERAGE** □**EXCEPTIONAL**	
Project performance is based on your average project grade. Possible performance levels range from receiving no points on any projects to receiving full points on all projects to date. A rating of "poor" indicates an average project score of 0%-45%. A rating of "average" indicates an average project score of 70%-79%. A rating of "exceptional" indicates an average project score of 90%-100%.	

Appendix N:
Distractor Form 4

PERFORMANCE FEEDBACK	
I. **STUDENT INFORMATION**	
STUDENT NAME	COURSE CODE
II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER	
III. PRIMARY PURPOSE - Provide information on student's class performance.	
IV. FEEDBACK	
1. EXAMS (45% of total grade)	
RATING: □ **1** □**3** □**5**	
Exam performance is based on your average exam score. Possible performance levels range from receiving no points on any exams to receiving full points on all exams to date. A rating of "poor" indicates an average exam score of 0%-45%. A rating of "average" indicates an average exam score of 70%-79%. A rating of "exceptional" indicates an average exam score of 90%-100%.	
2. QUIZZES (20% of total grade)	
RATING: □ **1** □**3** □**5**	
Quiz performance is based on participation. Possible performance levels range from zero quiz participation to participation in all quizzes. A rating of "poor" indicates no more than 40% of quizzes have been submitted. A rating of "average" indicates no more than 75% of quizzes have been submitted. A rating of "exceptional" indicates at least 90% of quizzes have been submitted.	
3. PROJECTS (35% of total grade)	
RATING: □ **1** □**3** □**5**	
Project performance is based on your average project grade. Possible performance levels range from receiving no points on any projects to receiving full points on all projects to date. A rating of "poor" indicates an average project score of 0%-45%. A rating of "average" indicates an average project score of 70%-79%. A rating of "exceptional" indicates an average project score of 90%-100%.	

Appendix O:
Distractor Form 5

PERFORMANCE FEEDBACK	
I. **STUDENT INFORMATION**	
STUDENT NAME	COURSE CODE
II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER	
III. PRIMARY PURPOSE - Provide information on student's class performance.	
IV. FEEDBACK	
1. TESTS (30% of total grade)	
RATING: □ **1** □**2** □**3** □**4** □**5**	
Test performance is based on your average test score. Possible performance levels range from receiving no points on any tests to receiving full points on all tests to date. A rating of "poor" indicates an average test score of 0%-45%. A rating of "average" indicates an average test score of 70%-79%. A rating of "exceptional" indicates an average test score of 90%-100%.	
2. ATTENDANCE (20% of total grade)	
RATING: □ **1** □**2** □**3** □**4** □**5**	
Attendance is based on coming to class and staying until class is dismissed, and is monitored by a sign-in sheet. Possible performance levels range from zero attendance to attendance of all classes. A rating of "poor" indicates no more than 40% of classes have been attended. A rating of "average" indicates no more than 75% of classes have been attended. A rating of "exceptional" indicates at least 90% of classes have been attended.	
3. FINAL RESEARCH PAPER (50% of total grade)	
RATING: □ **1** □**2** □**3** □**4** □**5**	
Research paper performance is based on one final research paper. Possible performance levels range from receiving no points on this paper to receiving full points this paper. A rating of "poor" indicates an average research paper score of 0%-45%. A rating of "average" indicates an average research paper score of 70%-79%. A rating of "exceptional" indicates an average research paper score of 90%-100%.	

Appendix P:
Distractor Form 6

PERFORMANCE FEEDBACK	
I. **STUDENT INFORMATION**	
STUDENT NAME	COURSE CODE
II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER	
III. PRIMARY PURPOSE - Provide information on student's class performance.	
IV. FEEDBACK	
1. EXAMS (forty five percent of total grade)	
RATING: ¢ 1 ¢2 ¢3 ¢4 ¢5	
Exam performance is based on your average exam score. Possible performance levels range from receiving no points on any exams to receiving full points on all exams to date. A rating of "poor" indicates an average exam score of 0%-45%. A rating of "average" indicates an average exam score of 70%-79%. A rating of "exceptional" indicates an average exam score of 90%-100%.	
2. QUIZZES (Twenty percent of total grade)	
RATING: ¢ 1 ¢2 ¢3 ¢4 ¢5	
Quizzes are an important part of your course grade. A rating of "poor" indicates that few or no quizzes have been submitted. A rating of "average" indicates that most quizzes have been submitted. A rating of "exceptional" indicates that mostly all quizzes have been submitted.	
3. PROJECTS (thirty five of total grade)	
RATING: ¢ 1 ¢2 ¢3 ¢4 ¢5	
Project performance is based on your average project grade. Possible performance levels range from receiving no points on any projects to receiving full points on all projects to date. A rating of "poor" indicates an average project score of 0%-45%. A rating of "average" indicates an average project score of 70%-79%. A rating of "exceptional" indicates an average project score of 90%-100%.	

Appendix Q:
Distractor Form 7

PERFORMANCE FEEDBACK	
I. **STUDENT INFORMATION**	
STUDENT NAME	COURSE CODE
II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER	
III. PRIMARY PURPOSE - Provide information on student's class performance.	
IV. FEEDBACK	
1. EXAMS (40% of total grade)	
RATING: □ **1**	□**3** □**5**
Exams are an important part of your course grade. Possible performance levels range from "poor" to "exceptional". A rating of "poor" indicates very low exam scores. A rating of "average" indicates passing exam scores. A rating of "exceptional" indicates high exam scores.	
2. QUIZZES (20% of total grade)	
RATING: □ **1**	□**3** □**5**
Quizzes are an important part of your course grade. A rating of "poor" indicates that few or no quizzes have been submitted. A rating of "average" indicates that most quizzes have been submitted. A rating of "exceptional" indicates that mostly all quizzes have been submitted.	
3. PROJECTS (35% of total grade)	
RATING: □ **1**	□**3** □**5**
Projects are an important part of your course grade. A rating of "poor" indicates very low project scores. A rating of "average" indicates passing project scores. A rating of "exceptional" indicates high project scores.	

Appendix R:
Distractor Form 8

PERFORMANCE FEEDBACK

I. **STUDENT INFORMATION**

STUDENT NAME	COURSE CODE

II. TYPE OF FEEDBACK: __ MID SEMESTER __ END OF SEMESTER

III. PRIMARY PURPOSE

- Provide information on student's class performance.

IV. FEEDBACK

1. EXAMS (Forty five percent of total grade)

RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL

- Exam performance is based on your average exam score.
- Possible performance levels range from receiving no points on any exams to receiving full points on all exams to date.
- A rating of "poor" indicates an average exam score of 0%-45%.
- A rating of "average" indicates an average exam score of 70%-79%.
- A rating of "exceptional" indicates an average exam score of 90%-100%.

2. QUIZZES (Twenty percent of total grade)

RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL

Quizzes are an important part of your course grade. A rating of "poor" indicates that few or no quizzes have been submitted. A rating of "average" indicates that most quizzes have been submitted. A rating of "exceptional" indicates that mostly all quizzes have been submitted.

3. PROJECTS (Thirty five percent of total grade)

RATING: ¢ POOR ¢BELOW AVERAGE ¢AVERAGE ¢ABOVE AVERAGE ¢EXCEPTIONAL

- Project performance is based on your average project grade.
- Possible performance levels range from receiving no points on any projects to receiving full points on all projects to date.
- A rating of "poor" indicates very low scores.
- A rating of "average" indicates average scores.
- A rating of "exceptional" indicates high scores.

Appendix S:
Scale Items

Please answer the following questions, with 1 = highly disagree and 5 = highly agree.

1. This form could provide performance feedback that is free of bias.

   1 2 3 4 5

2. This form could provide performance feedback that is based on accurate information.

   1 2 3 4 5

3. This form could provide feedback that is conveyed consistently.

   1 2 3 4 5

4. I would be satisfied with receiving this performance feedback form.

   1 2 3 4 5

5. I would be satisfied if my performance was evaluated using this feedback form.

   1 2 3 4 5

6. I am satisfied with the way this form presents my performance feedback.

   1 2 3 4 5

7. This form could provide an accurate evaluation of class performance.

   1 2 3 4 5

8. This form could reflect feedback that is based on true class performance.

   1 2 3 4 5

9. This form matches my own idea of class performance feedback.

   1 2 3 4 5

We would like to know why or your rationale for the responses you provided above. Please indicate the extent to which your responses were based on each of the factors below (1= not at all; 3= to a moderate extent; 5=to a significant extent)

1. The explanations of high, average and low performance are clear.

2. The rating options are clear.
3. The meaning of the rating options was clearly explained.
4. I would trust instructors to use this form appropriately.
5. The rating options are not clear enough.
6. The form is too busy with information.
7. This form seems unfair.

Please rank order the forms in order of preference, with 1 = most preferred and 4 = least preferred.

1. FORM A
2. FORM B
3. FORM C
4. FORM D

Appendix T:
Correlation Matrix of all Dependent and Mediating Variables

	Acceptance (V1)	Acceptance (V2)	Satisfaction (V1)	Satisfaction (V2)	Fairness (V1)	Fairness (V2)
Acceptance (V1)	___					
Acceptance (V2)	.734**	___				
Satisfaction (V1)	.843**	.638**	___			
Satisfaction (V2)	.649**	.841**	.681**	___		
Fairness (V1)	.640**	.523**	.591**	.437**	___	
Fairness (V2)	.436**	.676**	.383**	.572**	.742**	___

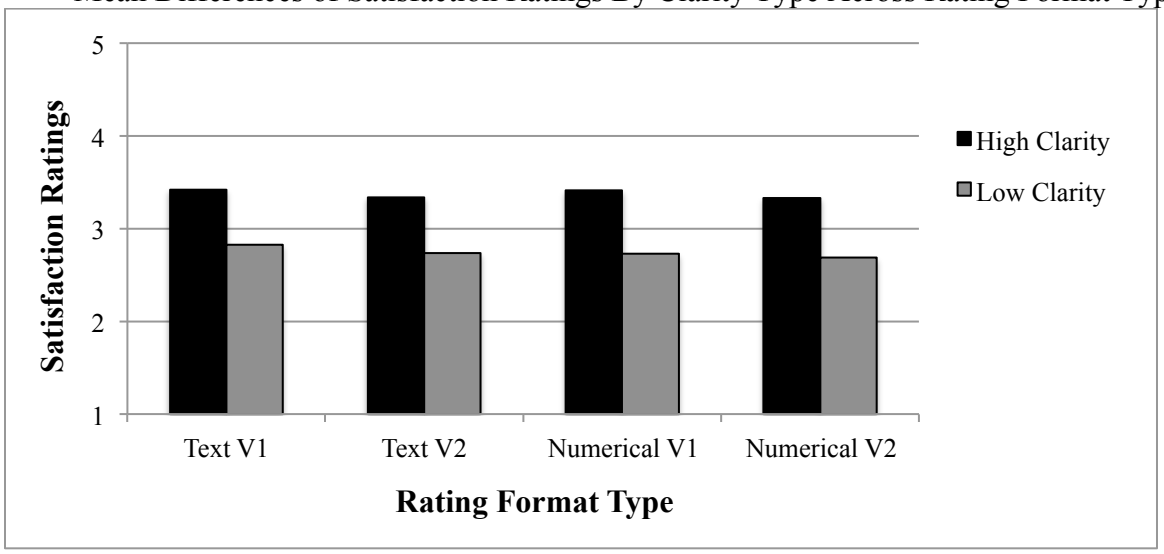
**$p < .001$.

Appendix U:
Mean of Responses Regarding Feedback Forms on Satisfaction, Acceptance, and Fairness Scales (with Standard Deviations in Parentheses)

Form Type	*n*	DV Satisfaction	DV Acceptance	Mediator Fairness
High Clarity and Text (V1)	272	3.43 (.95)	3.41 (.86)	3.68 (.79)
High Clarity and Text (V2)	272	3.41 (.92)	3.37 (.88)	3.68 (.79)
High Clarity and Numerical (V1)	272	3.34 (.96)	3.36 (.87)	3.69 (.75)
High Clarity and Numerical (V2)	272	3.33 (.99)	3.39 (.88)	3.69 (.79)
Low Clarity and Text (V1)	272	2.83 (1.00)	2.88 (.97)	3.28 (.86)
Low Clarity and Text (V2)	272	2.73 (.99)	2.84 (.95)	3.25 (.85)
Low Clarity and Numerical (V1)	272	2.74 (1.00)	2.83 (.95)	3.27 (.85)
Low Clarity and Numerical (V2)	272	2.69 (.97)	2.78 (.91)	3.20 (.88)

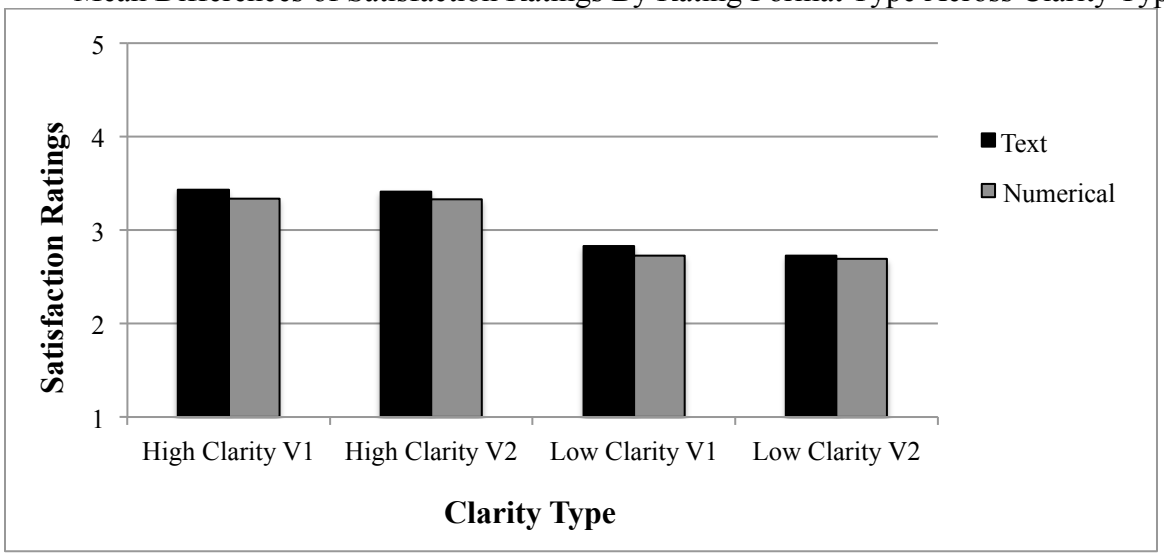
Key. V1 = Version 1, V2 = Version 2.
**Note: Maximum response possible was 5.00 (Strongly Agree)*

Appendix V:
Mean Differences of Satisfaction Ratings By Clarity Type Across Rating Format Type



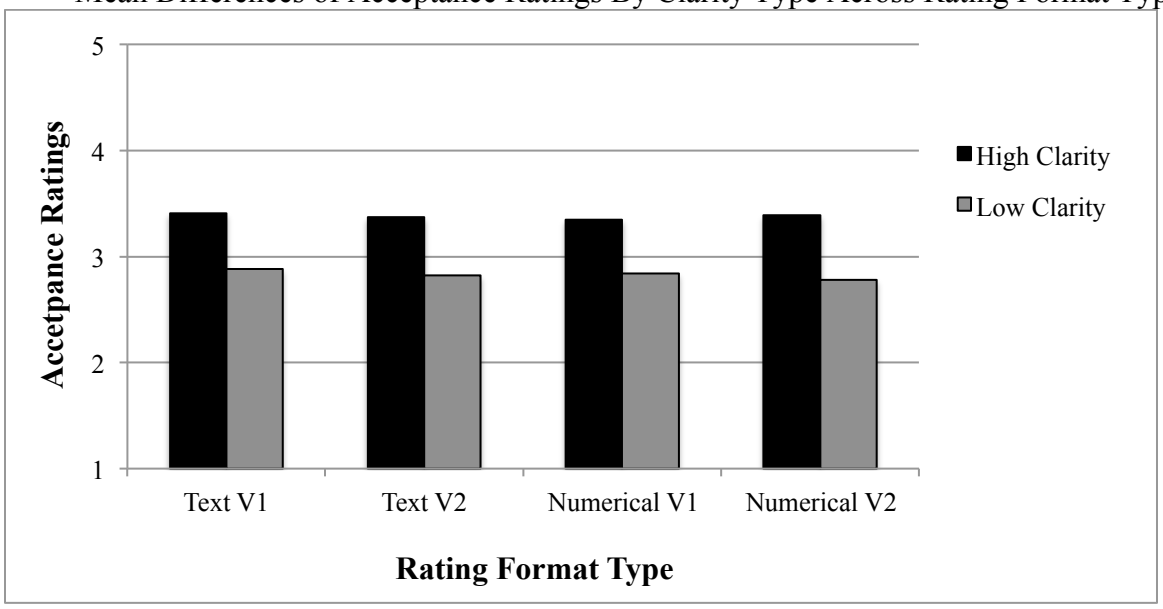
Appendix W:

Mean Differences of Satisfaction Ratings By Rating Format Type Across Clarity Type

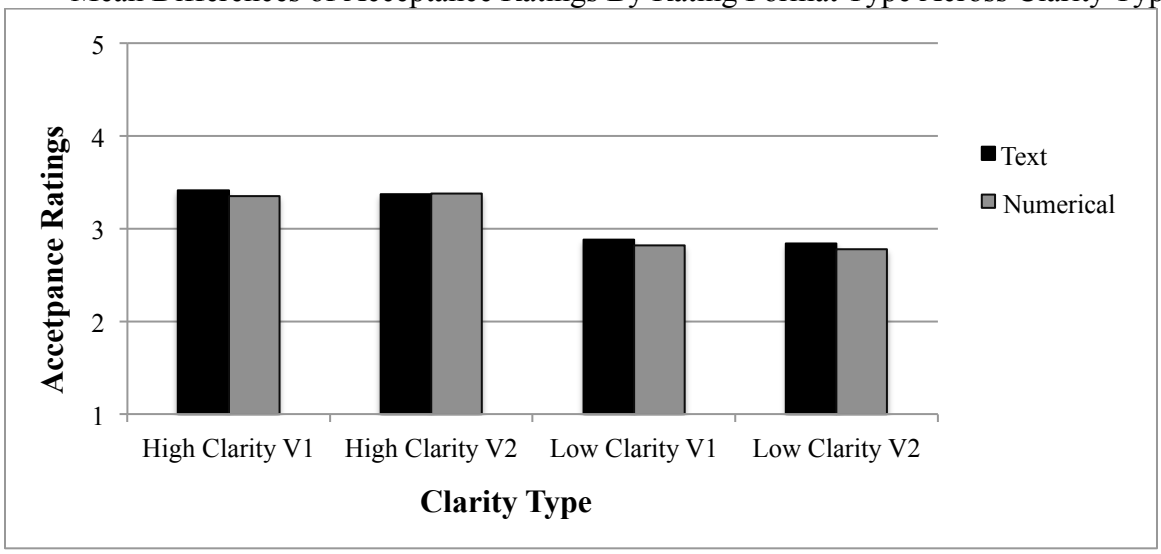


Appendix X:

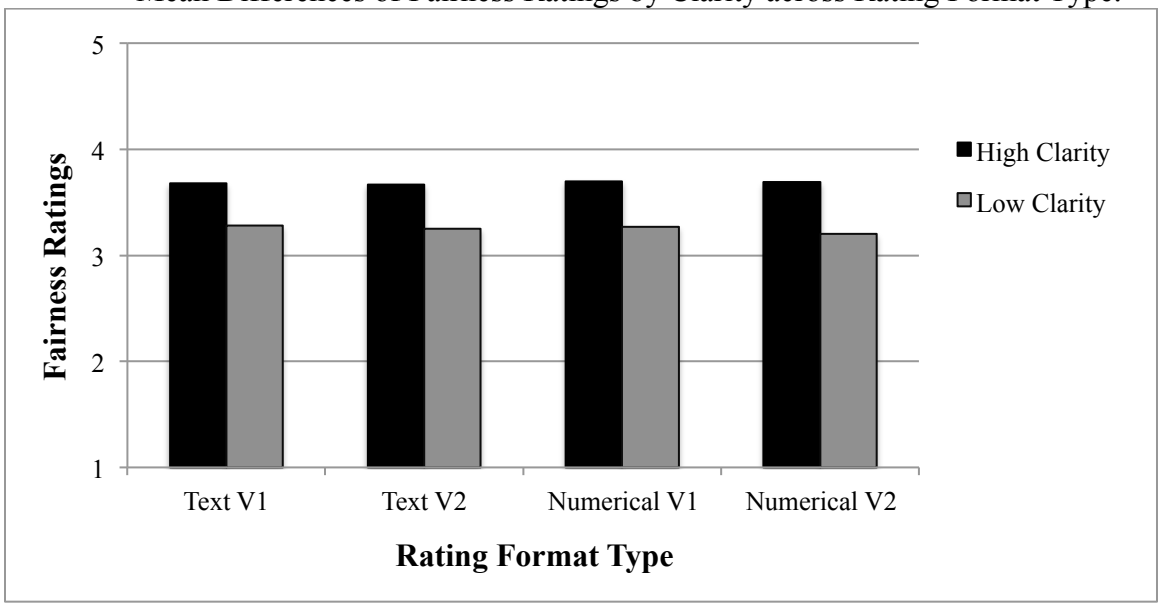
Mean Differences of Acceptance Ratings By Clarity Type Across Rating Format Type



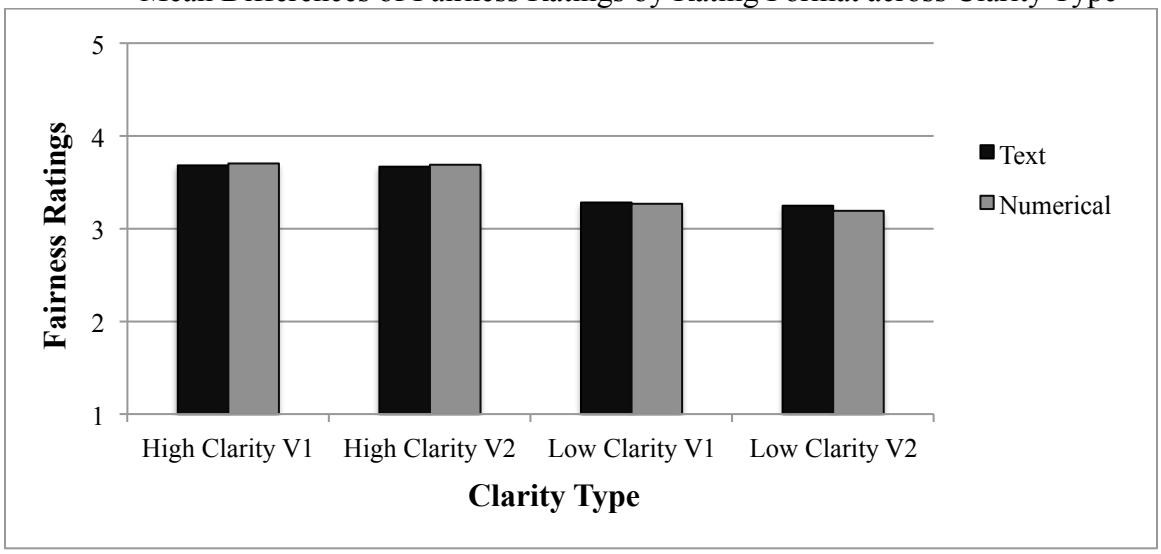
Appendix Y:
Mean Differences of Acceptance Ratings By Rating Format Type Across Clarity Type.



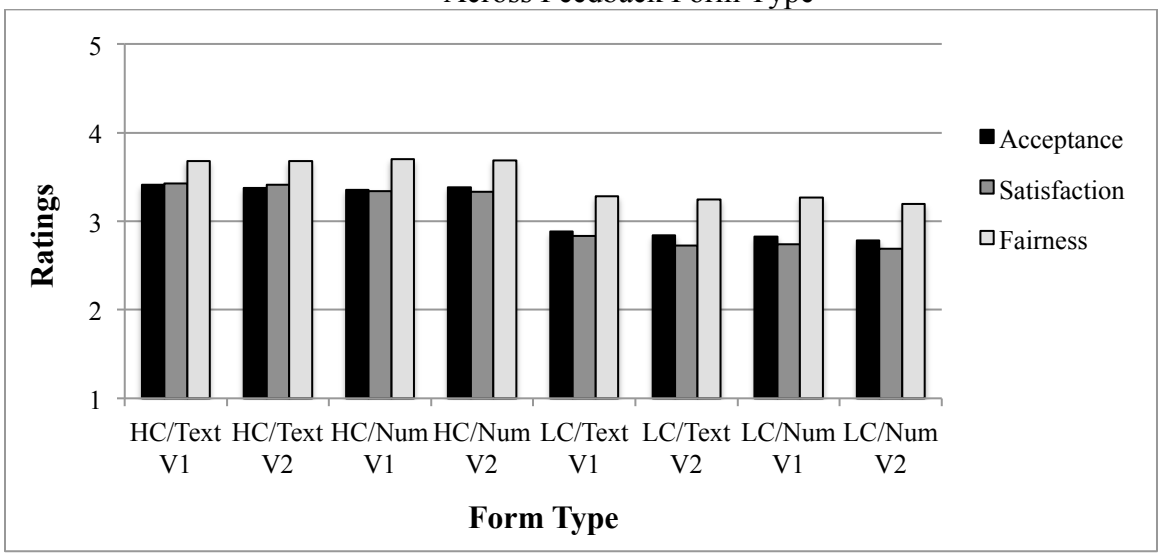
Appendix Z:
Mean Differences of Fairness Ratings by Clarity across Rating Format Type.



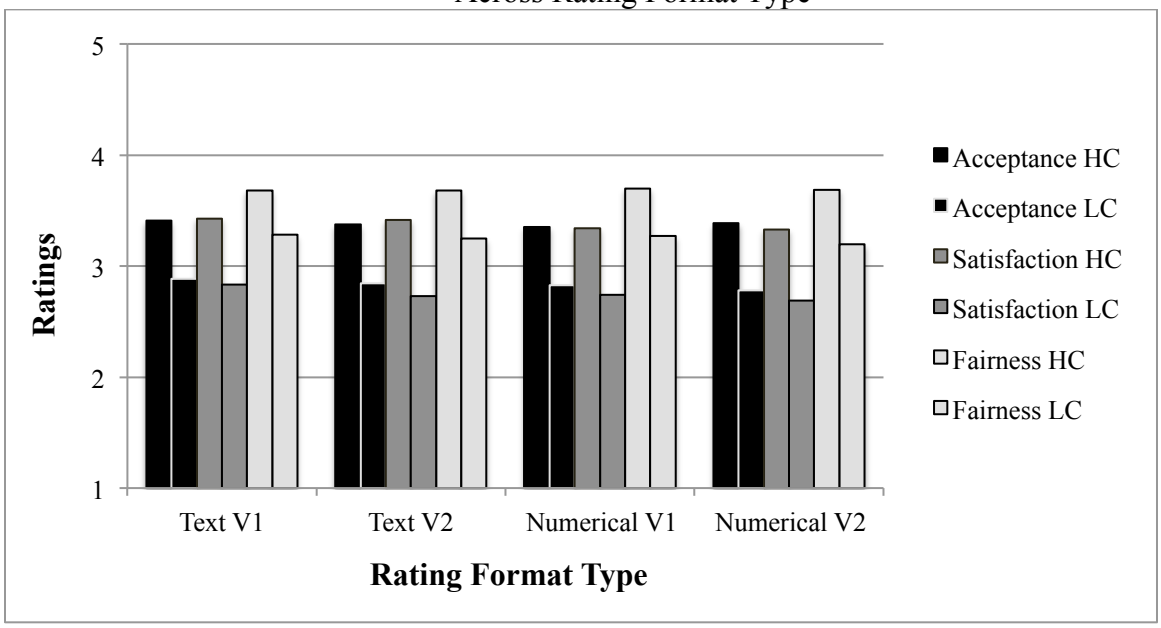
Appendix AA:
Mean Differences of Fairness Ratings by Rating Format across Clarity Type



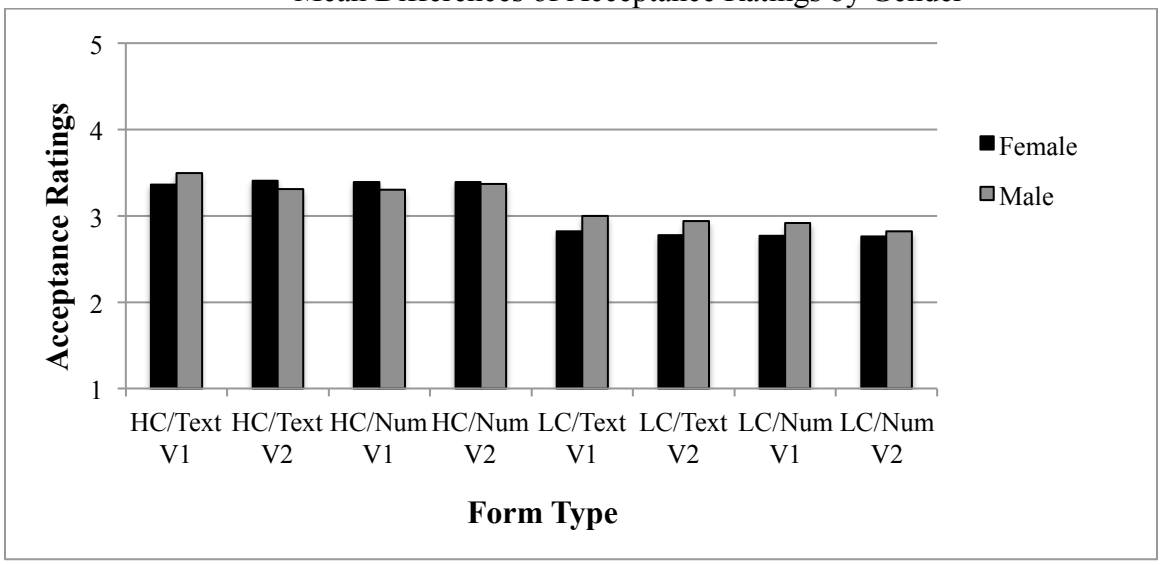
Appendix AB:
Mean Differences of Satisfaction, Acceptance, and Fairness Ratings
Across Feedback Form Type



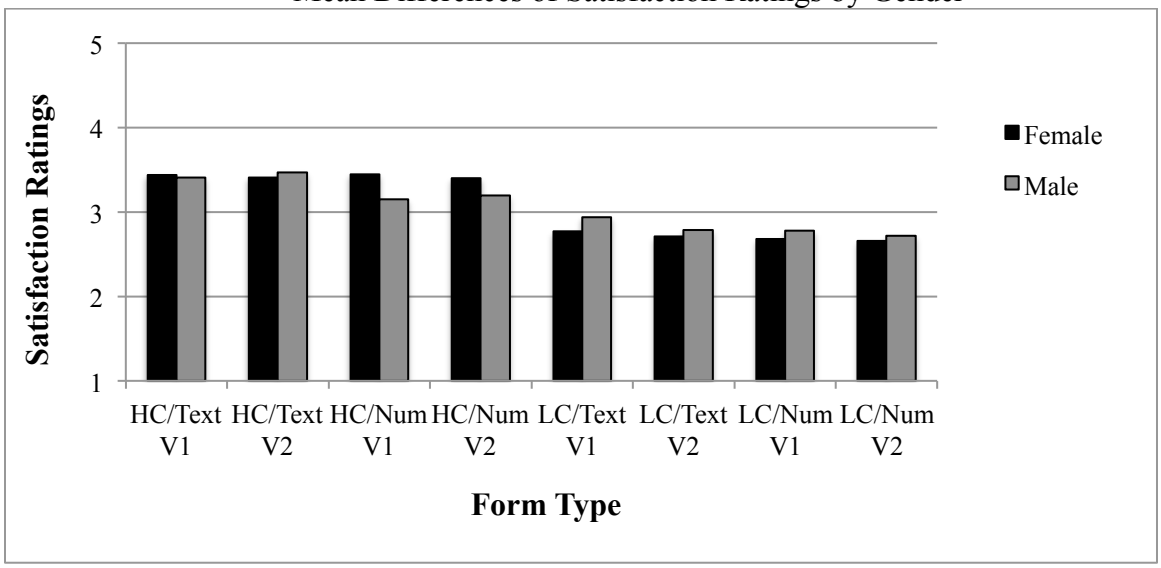
Appendix AC:
Mean Differences of Satisfaction, Acceptance, and Fairness Ratings By Clarity Across Rating Format Type



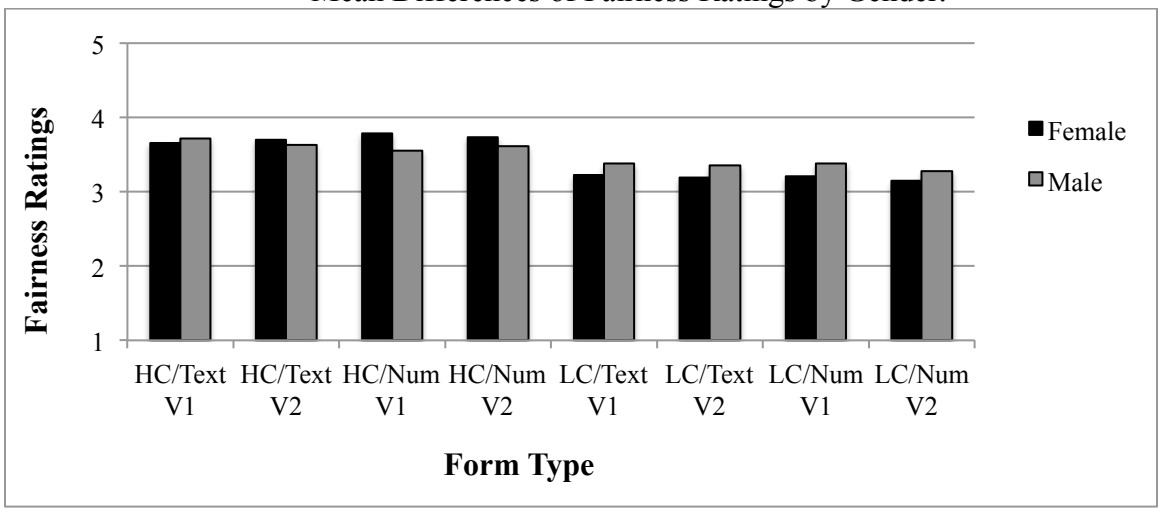
Appendix AD:
Mean Differences of Acceptance Ratings by Gender



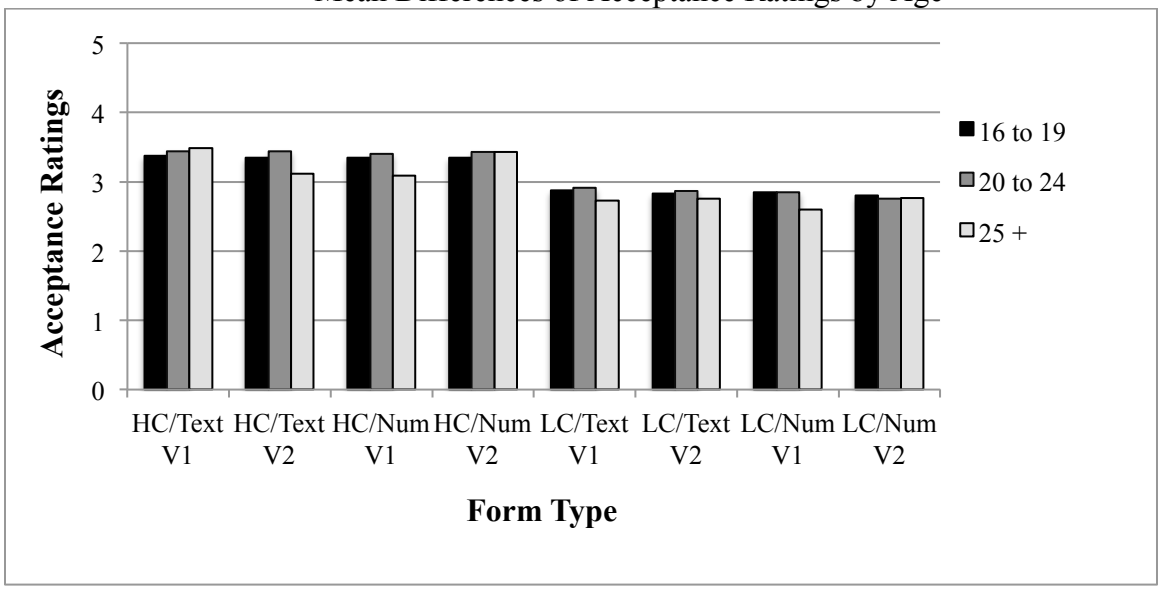
Appendix AE:
Mean Differences of Satisfaction Ratings by Gender



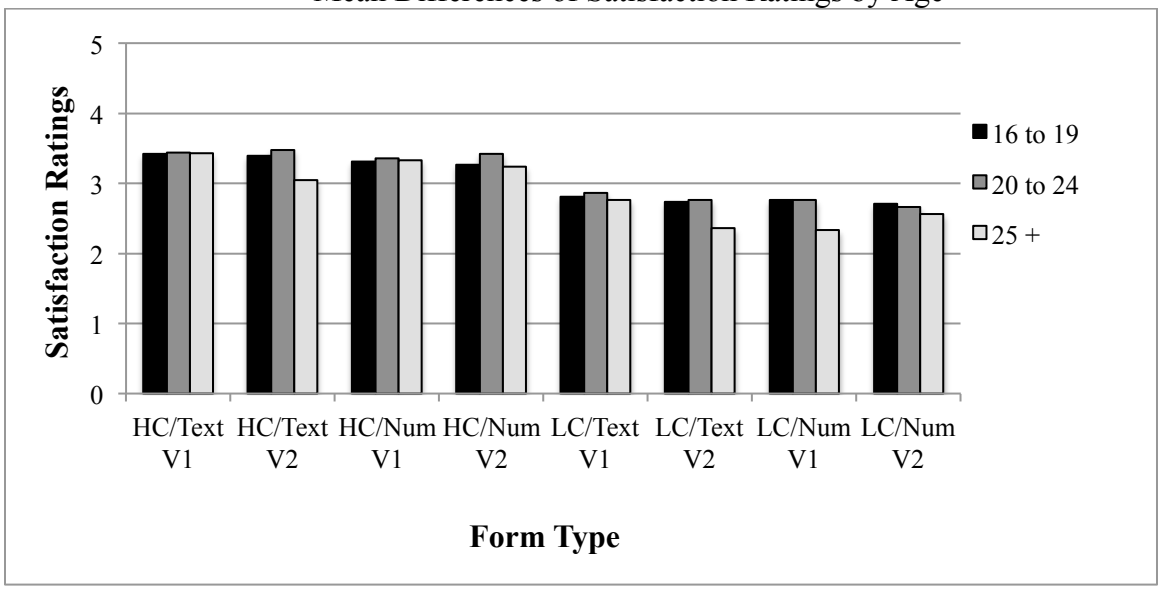
Appendix AF:
Mean Differences of Fairness Ratings by Gender.



Appendix AG:
Mean Differences of Acceptance Ratings by Age



Appendix AH:
Mean Differences of Satisfaction Ratings by Age



Appendix AI:
Mean Differences of Fairness Ratings by Age